GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 12:03:10 ; Search time 4311 Seconds
                                (without alignments)
                                10688.319 Million cell updates/sec

Title: US-09-891-138A-1
Perfect score: 1543
Sequence: 1 gctcctggcagagttttctg..........tgcctaaataaatcaatata 1543

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   EST:*
               1:   em_estba:*
               2:   em_esthum:*
               3:   em_estin:*
               4:   em_estmu:*
               5:   em_estov:*
               6:   em_estpl:*
               7:   em_estro:*
               8:   em_htc:*
               9:   gb_est1:*
               10: gb_est2:*
               11: gb_htc:*
               12: gb_est3:*
               13: gb_est4:*
               14: gb_est5:*
               15: em_estfun:*
               16: em_estom:*
               17: em_gss_hum:*
               18: em_gss_inv:*
               19: em_gss_pln:*
               20: em_gss_vrt:*
               21: em_gss_fun:*
               22: em_gss_mam:*
               23: em_gss_mus:*
               24: em_gss_pro:*
               25: em_gss_rod:*
               26: em_gss_phg:*
               27: em_gss_vrl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1484.8	96.2	1585	11	AK080866	AK080866 Mus muscu
	2	560	36.3	683	10	BB323771	BB323771 BB323771
	3	516.4	33.5	556	13	BX527630	BX527630 BX527630
	4	495.8	32.1	520	9	AI663305	AI663305 uk27c10.y
	5	455	29.5	469	10	BB744515	BB744515 BB744515
	6	438	28.4	458	10	BB746222	BB746222 BB746222
	7	414	26.8	428	10	BB738743	BB738743 BB738743
	8	403.8	26.2	422	10	BB847918	BB847918 BB847918
	9	388.4	25.2	420	10	BB864882	BB864882 BB864882
	10	384.8	24.9	426	10	BB778587	BB778587 BB778587
	11	380.4	24.7	396	10	BB739482	BB739482 BB739482
c	12	363.8	23.6	367	9	AI649254	AI649254 uk27c10.x
	13	357.6	23.2	636	10	BB645274	BB645274 BB645274
	14	354.2	23.0	416	10	BB846608	BB846608 BB846608
	15	350.6	22.7	408	13	BY368584	BY368584 BY368584
	16	346.6	22.5	868	14	CD246161	CD246161 AGENCOURT
	17	319.4	20.7	497	29	CE610929	CE610929 tigr-gss-
	18	309.8	20.1	327	10	BB220946	BB220946 BB220946
	19	296	19.2	877	12	BG402029	BG402029 602466748
	20	294.6	19.1	333	10	BB254869	BB254869 BB254869
	21	283.2	18.4	323	10	BB220888	BB220888 BB220888
	22	279	18.1	316	10	BB225749	BB225749 BB225749
	23	276.2	17.9	501	13	BX281458	BX281458 BX281458
	24	275.2	17.8	326	10	BB500452	BB500452 BB500452
	25	269.4	17.5	770	13	BU141159	BU141159 603137524
	26	261.4	16.9	285	10	BB327439	BB327439 BB327439
c	27	255.8	16.6	633	29	AG083174	AG083174 Pan trogl
	28	253.4	16.4	314	10	BB498575	BB498575 BB498575
	29	241	15.6	310	10	BB221521	BB221521 BB221521
	30	218.4	14.2	805	13	BU291924	BU291924 603606116
	31	216.8	14.1	366	13	BY005778	BY005778 BY005778
	32	215.4	14.0	282	10	BB215653	BB215653 BB215653
	33	214	13.9	312	10	BB498898	BB498898 BB498898
	34	202	13.1	1101	29	CNS04VV3	AL309576 Tetraodon
	35	199.8	12.9	226	10	BB213317	BB213317 BB213317
c	36	192.8	12.5	956	29	CNS028Y4	AL186565 Tetraodon
	37	189.6	12.3	655	13	BU373390	BU373390 603811071
c	38	181	11.7	294	9	AW112068	AW112068 MC15648 m
c	39	158.8	10.3	639	10	AW612141	AW612141 hg94h07.x
	40	151.2	9.8	657	13	BU352057	BU352057 603527490
	41	151	9.8	1026	29	CNS051MY	AL317059 Tetraodon
c	42	144.2	9.3	589	10	BF196066	BF196066 hr81f02.x
	43	141	9.1	1022	29	CNS04W90	AL310077 Tetraodon
c	44	139.8	9.1	582	10	BE221739	BE221739 hr58c09.x
	45	137.4	8.9	139	9	AI021184	AI021184 ub02f12.r

RESULT 1
AK080866
```
LOCUS       AK080866                1585 bp    mRNA    linear   HTC 20-SEP-2003
DEFINITION  Mus musculus 4 days neonate male adipose cDNA, RIKEN full-length
            enriched library, clone:B430012O21 product:G-PROTEIN COUPLED
            RECEPTOR GPR91, full insert sequence.
ACCESSION   AK080866
VERSION     AK080866.1  GI:26099527
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 1585)
```

AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
TITLE   Direct Submission
JOURNAL   Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
COMMENT   cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
FEATURES   Location/Qualifiers
    source   1. .1585
                    /organism="Mus musculus"
                    /mol_type="mRNA"
                    /strain="C57BL/6J"
                    /db_xref="FANTOM_DB:B430012O21"
                    /db_xref="MGI:2411980"
                    /db_xref="taxon:10090"
                    /clone="B430012O21"
                    /sex="male"
                    /tissue_type="adipose"
                    /clone_lib="RIKEN full-length enriched mouse cDNA library"
                    /dev_stage="4 days neonate"
    misc_feature   69. .1025
                    /note="G-PROTEIN COUPLED RECEPTOR GPR91 (SPTR|Q99MT6,
                    evidence: FASTY, 94.3%ID, 100%length, match=954)
                    putative"
    polyA_signal   1558. .1563
                    /note="putative"
    polyA_site   1585
                    /note="putative"
ORIGIN

    Query Match   96.2%; Score 1484.8; DB 11; Length 1585;
    Best Local Similarity   98.4%; Pred. No. 0;
    Matches 1521; Conservative   0; Mismatches   22; Indels   3; Gaps   2;

Qy   1 GCTCCTGGCAGAGTTTTCTGTCGAGACAGAAGCCGACAGCAGAATGGCACAGAATTTATC 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db     26 GCTCCTGGCAGAGTTTTCTGTCGAGACAGAAGCCGACAGCAGAATGGCACAGAATTTATC 85

Qy     61 TTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTA 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db     86 TTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGGATTTTA 145

Qy    121 TGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGTGTTCGGCTACCTCTT 180
          |||||||||||||||||||||||||||||||||||||||||||| |||||| || |||||| ||
Db    146 TGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGGGGTGTTTGGGTACCTGTT 205

Qy    181 CTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCTTTCCATCTCTGACTT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    206 CTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCTTTCCATCTCTGACTT 265

Qy    241 TGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAATGATAAGGGGACCTA 300
          ||||||||| ||||||||||||||||||||||||||||| |||||||||||||||||||||
Db    266 TGCTTTCCTGGGCACCCTTCCCATCCTGATAAAGAGTTTTGCCAATGATAAGGGGACCTA 325

Qy    301 TGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAACCTCTACACCAGCAT 360
          ||||||||||| || |||||||||||||||||| ||||| |||||||||| || ||||||||
Db    326 TGGAGATGTTCTTTGGATAAGCAACCGATATGGGCTTAACACCAACCTTTAAACCAGCAT 385

Qy    361 CCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGCTCATGAA--GTACCCTTTCCGA 418
          | | ||| ||| ||||||||||||||||||||||||||||||||||      || ||||||||
Db    386 CTTTTTCTTCATTTTCATTAGCATGGACCGATATCTGCTCATGAAAGTACCCTTTTCCGA 445

Qy    419 GAACAC-TTTCTACAAAAGAAGGAATTTGCCATTTTAATCTCGCTGGCTGTCTGGGCCTT 477
          ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    446 GAACACTTTTCTACAAAAGAAGGAATTTGCCATTTTAATCTCGCTGGCTGTCTGGGCCTT 505

Qy    478 AGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATTCTGTCCCAAAAGAAGAGGG 537
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    506 AGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATTCTGTCCCAAAAGAAGAGGG 565

Qy    538 CAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAACACAATCTCATTTACAGCCT 597
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    566 CAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAACACAATCTCATTTACAGCCT 625

Qy    598 CTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTGTGATGTGCTTCTTCTACTACAA 657
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    626 CTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTGTGATGTGCTTCTTCTACTACAA 685

Qy    658 GATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAACTGCCCTGCCACTGGACAA 717
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    686 GATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAACTGCCCTGCCACTGGACAA 745

Qy    718 ACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTATACTCTTCACACCCTATCA 777
          ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    746 ACCCCAACGCCTGGTGGTCCTGGCAGTTGTGATCTTCTCTATACTCTTCACACCCTATCA 805

Qy    778 TATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTTGGCCACAAGGATGTACACA 837
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    806 TATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTTGGCCACAAGGATGTACACA 865

Qy    838 GAAGGCCATCAAATCTATATACACACTGACACGGCCTCTGGCCTTTCTGAACAGTGCCAT 897
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    866 GAAGGCCATCAAATCTATATACACACTGACACGGCCTCTGGCCTTTCTGAACAGTGCCAT 925
```

```
Qy    898  CAATCCCATCTTCTACTTCCTCATGGGAGACCATTACAGAGAGATGCTGATTAGTAAGTT  957
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    926  CAATCCCATCTTCTACTTCCTCATGGGAGACCATTACAGAGAGATGCTGATTAGTAAGTT  985

Qy    958  CAGACAATACTTCAAGTCCCTTACATCCTTCAGGACATGAGCTGCTGGATGCAGGTCTTC  1017
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    986  CAGACAATACTTCAAGTCCCTTACATCCTTCAGGACATGAGCTGCTGGATGCAGGTCTTC  1045

Qy    1018 ACTCAGCCAAAATGAGACACTTGATAAACAGTGCTGTGCAGTTGAGTTTTAACTAAGTAA  1077
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1046 ACTCAGCCAAAATGAGACACTTGATAAACAGTGCTGTGCAGTTGAGTTTTAACTAAGTAA  1105

Qy    1078 ACCACCATTTCTAGGCTTTAGCTTTCCACCATCCTCCAACCCCCAGGGCTGGAGTACAAG  1137
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1106 ACCACCATTTCTAGGCTTTAGCTTTCCACCATCCTCCAACCCCCAGGGCTGGAGTACAAG  1165

Qy    1138 CTGGGTCCACATGAATCAGAAGGCAGCTCTCTGTTCTGATTTTAGGTTATACCCAGAGTA  1197
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1166 CTGGGTCCACATGAATCAGAAGGCAGCTCTCTGTTCTGATTTTAGGTTATACCCAGAGTA  1225

Qy    1198 TGGAAAAAATAAGGCATGAGAAAGCATTGACATCTTCACTTAAGAACTGAACAAAAGAGA  1257
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1226 TGGAAAAAATAAGGCATGAGAAAGCATTGACATCTTCACTTAAGAACTGAACAAAAGAGA  1285

Qy    1258 ACAAATATTGTCAATGTTTGGACACTTAGGATCTGAAATCTTGGAAATTTTAAGACCTCT  1317
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1286 ACAAATATTGTCAATGTTTGGACACTTAGGATCTGAAATCTTGGAAATTTTAAGACCTCT  1345

Qy    1318 TTTTCTATCAGTGTAAAAGGAATACAAGATAGCTAGTTGCAAATGCTGAATGCATTTCAT  1377
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1346 TTTTCTATCAGTGTAAAAGGAATACAAGATAGCTAGTTGCAAATGCTGAATGCATTTCAT  1405

Qy    1378 CATTGGTCAGGTCGATAAGCGTGTTTCTGAAATAGTCTTATTTTTATTCTTGTAATATTA  1437
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1406 CATTGGTCAGGTCGATAAGCGTGTTTCTGAAATAGTCTTATTTTTATTCTTGTAATATTA  1465

Qy    1438 AAATTTATGTGAAAAATGAATATAATTCAATGTACAACATTAGATTTTCTATTTGAAAAT  1497
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1466 AAATTTATGTGAAAAATGAATATAATTCAATGTACAACATTAGATTTTCTATTTGAAAAT  1525

Qy    1498 TATATTTCTTGAAAAAATAACTGCTGTGCCTAAATAAATCAATATA  1543
           ||||||||||||||||||||||||||||||||||||||||||||||
Db    1526 TATATTTCTTGAAAAAATAACTGCTGTGCCTAAATAAATCAATATA  1571
```

```
RESULT 2
BB323771
LOCUS         BB323771              683 bp    mRNA    linear    EST 31-AUG-2001
DEFINITION    BB323771 RIKEN full-length enriched, 4 days neonate male adipose
              Mus musculus cDNA clone B430012O21 3', mRNA sequence.
ACCESSION     BB323771
VERSION       BB323771.2  GI:15411432
KEYWORDS      EST.
SOURCE        Mus musculus (house mouse)
  ORGANISM    Mus musculus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
                 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE     1  (bases 1 to 683)
   AUTHORS      Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
                Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
                Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
                Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
                Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
                Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
                Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
   TITLE        RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
   JOURNAL      Unpublished (2001)
   COMMENT      On Jul 11, 2000 this sequence version replaced gi:9032085.
                Contact: Yoshihide Hayashizaki
                Laboratory for Genome Exploration Research Group, RIKEN Genomic
                Sciences Center(GSC), Yokohama Institute
                The Institute of Physical and Chemical Research (RIKEN)
                1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
                Tel: 81-45-503-9222
                Fax: 81-45-503-9216
                Email: genome-res@gsc.riken.go.jp,
                URL:http://genome.gsc.riken.go.jp/
                Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
                Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
                 Normalization and subtraction of cap-trapper-selected cDNAs to
                prepare full-length cDNA libraries for rapid discovery of new
                genes. Genome Res. . 10 (10), 1617-1630 (2000)
                 wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
                Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
                Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
                and Hayashizaki,Y.
                 RIKEN integrated sequence analysis (RISA) system--384-format
                sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
                10 (11), 1757-1771 (2000)
                 Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
                Sugahara,Y. and Hayashizaki,Y.
                 Computer-based methods for the mouse full-length cDNA
                encyclopedia: real-time sequence clustering for construction of a
                nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
                 Yamanaka,I., Kiyosawa,H., Kondo,S., Saito,T., Shinagawa,A.,
                Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K.,
                Arakawa,T., Ishii,Y. and Hayashizaki,Y.
                 Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct.
                Func. Genomics 2 pre, L72-L86 (2001)
                 Please visit our web site (http://genome.gsc.riken.go.jp/) for
                further details.
                cDNA library was prepared and sequenced in Mouse Genome
                Encyclopedia Project of Genome Exploration Research Group in Riken
                Genomic Sciences Center and Genome Science Laboratory in RIKEN.
                Division of Experimental Animal Research in Riken contributed to
                prepare mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .683
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="B430012021"
                     /sex="male"
```

/tissue_type="adipose"
/dev_stage="4 days neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 4 days neonate
male adipose"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 10.0 and subtraction to Rot = 229.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I."

ORIGIN

```
Query Match            36.3%;   Score 560;   DB 10;   Length 683;
 Best Local Similarity  98.7%;   Pred. No. 4.2e-124;
   Matches  596;  Conservative    0;  Mismatches    5;  Indels    3;  Gaps    3;

Qy        943  GCTGATTAGTAAGTTCAGAC-AATACTTCAAG-TCCCTTACATCCTTC-AGGACATGAGC  999
               |||||||||||||||||||| | ||||||||||| |||||||||||||| ||||||| ||
Db         66  GCTGATTAGTAAGTTCAGCCAAATACTTCAAGTTCCCTTACATCCTTCAAGGACATAAGT  125

Qy       1000  TGCTGGATGCAGGTCTTCACTCAGCCAAAATGAGACACTTGATAAACAGTGCTGTGCAGT  1059
               |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        126  TGCTGGATGCAGGTTTTCACTCAGCCAAAATGAGACACTTGATAAACAGTGCTGTGCAGT  185

Qy       1060  TGAGTTTTAACTAAGTAAACCACCATTTCTAGGCTTTAGCTTTCCACCATCCTCCAACCC  1119
               ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186  TGAGTTTTAATTAAGTAAACCACCATTTCTAGGCTTTAGCTTTCCACCATCCTCCAACCC  245

Qy       1120  CCAGGGCTGGAGTACAAGCTGGGTCCACATGAATCAGAAGGCAGCTCTCTGTTCTGATTT  1179
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246  CCAGGGCTGGAGTACAAGCTGGGTCCACATGAATCAGAAGGCAGCTCTCTGTTCTGATTT  305

Qy       1180  TAGGTTATACCCAGAGTATGGAAAAAATAAGGCATGAGAAAGCATTGACATCTTCACTTA  1239
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306  TAGGTTATACCCAGAGTATGGAAAAAATAAGGCATGAGAAAGCATTGACATCTTCACTTA  365

Qy       1240  AGAACTGAACAAAAGAGAACAAATATTGTCAATGTTTGGACACTTAGGATCTGAAATCTT  1299
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366  AGAACTGAACAAAAGAGAACAAATATTGTCAATGTTTGGACACTTAGGATCTGAAATCTT  425

Qy       1300  GGAAATTTTAAGACCTCTTTTTCTATCAGTGTAAAAGGAATACAAGATAGCTAGTTGCAA  1359
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426  GGAAATTTTAAGACCTCTTTTTCTATCAGTGTAAAAGGAATACAAGATAGCTAGTTGCAA  485

Qy       1360  ATGCTGAATGCATTTCATCATTGGTCAGGTCGATAAGCGTGTTTCTGAAATAGTCTTATT  1419
```

```
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     486 ATGCTGAATGCATTTCATCATTGGTCAGGTCGATAAGCGTGTTTCTGAAATAGTCTTATT 545

Qy     1420 TTTATTCTTGTAATATTAAAATTTATGTGAAAAATGAATATAATTCAATGTACAACATTA 1479
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     546 TTTATTCTTGTAATATTAAAATTTATGTGAAAAATGAATATAATTCAATGTACAACATTA 605

Qy     1480 GATTTTCTATTTGAAAATTATATTTCTTGAAAAAATAACTGCTGTGCCTAAATAAATCAA 1539
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     606 GATTTTCTATTTGAAAATTATATTTCTTGAAAAAATAACTGCTGTGCCTAAATAAATCAA 665

Qy     1540 TATA 1543
            ||||
Db     666 TATA 669
```

RESULT 3
BX527630

```
LOCUS       BX527630                 556 bp    mRNA    linear   EST 27-JUN-2003
DEFINITION  BX527630 Sugano mouse kidney mkia Mus musculus cDNA clone
            IMAGp998B194840 ; IMAGE:1970226, mRNA sequence.
ACCESSION   BX527630
VERSION     BX527630.1  GI:32297360
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 556)
  AUTHORS   Heil,O., Ebert,L., Neubert,P., Peters,M., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Mouse UnigeneSet - RZPD2
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
            RZPD; IMAGp998B194840.
            RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
            Mouse UnigeneSet - RZPD2 (RZPDLIB No.981)
            http://www.rzpd.de/CloneCards/cgi-
            bin/showLib.pl.cgi/response?libNo=981 Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 101
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available royalty-free from RZPD;
            contact RZPD (clone@rzpd.de) for further information. Seq primer:
            sugF, Primer sequence: CTTCTGCTCTAAAAGCTGCG.
FEATURES             Location/Qualifiers
     source          1. .556
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGp998B194840 ; IMAGE:1970226"
                     /sex="female"
```

/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse kidney mkia"
/note="Organ: kidney; Vector: pME18S-FL3; Site_1: DraIII
(CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCCTTTTTTTTTTTTTTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTTGGCCTACTGG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end
primer CGACCTGCAGCTCGAGCACA."

ORIGIN

```
Query Match              33.5%;   Score 516.4;  DB 13;  Length 556;
Best Local Similarity    99.8%;   Pred. No. 1.3e-113;
Matches   517;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;


Qy        1 GCTCCTGGCAGAGTTTTCTGTCGAGACAGAAGCCGACAGCAGAATGGCACAGAATTTATC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       39 GCTCCTGGCAGAGTTTTCTGTCGAGACAGAAGCCGACAGCAGAATGGCACAGAATTTATC 98


Qy       61 TTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       99 TTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTA 158


Qy      121 TGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGTGTTCGGCTACCTCTT 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db      159 TGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGTGTTTGGCTACCTCTT 218


Qy      181 CTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCTTTCCATCTCTGACTT 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      219 CTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCTTTCCATCTCTGACTT 278


Qy      241 TGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAATGATAAGGGGACCTA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      279 TGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAATGATAAGGGGACCTA 338


Qy      301 TGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAACCTCTACACCAGCAT 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      339 TGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAACCTCTACACCAGCAT 398


Qy      361 CCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGCTCATGAAGTACCCTTTCCGAGA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      399 CCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGCTCATGAAGTACCCTTTCCGAGA 458


Qy      421 ACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCTCGCTGGCTGTCTGGGCCTTAGT 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      459 ACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCTCGCTGGCTGTCTGGGCCTTAGT 518


Qy      481 GACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
            ||||||||||||||||||||||||||||||||||||||
Db      519 GACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 556
```

RESULT 4
AI663305
```
LOCUS       AI663305                 520 bp    mRNA    linear   EST 10-MAY-1999
DEFINITION  uk27c10.y1 Sugano mouse kidney mkia Mus musculus cDNA clone
            IMAGE:1970226 5' similar to SW:P2YR_RAT P49651 P2Y PURINOCEPTOR 1
            ;, mRNA sequence.
ACCESSION   AI663305
VERSION     AI663305.1  GI:4766888
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 520)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
            Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
            Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Other_ESTs: uk27c10.x1
            Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:986966
            Seq primer: custom primer used
            High quality sequence stop: 490.
FEATURES             Location/Qualifiers
     source          1. .520
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1970226"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="Sugano mouse kidney mkia"
                     /note="Organ: kidney; Vector: pME18S-FL3; Site_1: DraIII
                     (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCCTTTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a  DraIII adaptor [TGTTGGCCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
                     be used to isolate the cDNA insert.  Size selection was
                     performed to exclude fragments <1.5kb.  Library
                     constructed by Dr. Sumio Sugano (University of Tokyo
                     Institute of Medical Science).  Custom primers  for
```

ORIGIN

```
    Query Match              32.1%;   Score 495.8;  DB 9;  Length 520;
    Best Local Similarity    98.6%;   Pred. No. 1.2e-108;
    Matches  500;  Conservative    0;  Mismatches   7;  Indels   0;  Gaps   0;
```

```
Qy      1 GCTCCTGGCAGAGTTTTCTGTCGAGACAGAAGCCGACAGCAGAATGGCACAGAATTTATC 60
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db     14 GCTCCTGGCAGAGTTTTCTGTCGAGACAGAAGCCGACAGCTGAATGGCACAGAATTTATC 73

Qy     61 TTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTA 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     74 TTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTA 133

Qy    121 TGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGTGTTCGGCTACCTCTT 180
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    134 TGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGTGTTTGGCTACCTCTT 193

Qy    181 CTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCTTTCCATCTCTGACTT 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    194 CTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCTTTCCATCTCTGACTT 253

Qy    241 TGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAATGATAAGGGGACCTA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    254 TGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAATGATAAGGGGACCTA 313

Qy    301 TGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAACCTCTACACCAGCAT 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    314 TGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAACCTCTACACCAGCAT 373

Qy    361 CCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGCTCATGAAGTACCCTTTCCGAGA 420
          ||||| ||||||| ||||||  ||||||||||||||||||||||||||||| |||||||
Db    374 GCTCTTGCTCACTGTCATTATCATGGACCGATATCTGCTCATGAAGTACCCTGTCCGAGA 433

Qy    421 ACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCTCGCTGGCTGTCTGGGCCTTAGT 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    434 ACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCTCGCTGGCTGTCTGGGCCTTAGT 493

Qy    481 GACCTTAGAAGTTCTACCCATGCTCAC 507
          |||||||||||||||||||||||||||
Db    494 GACCTTAGAAGTTCTACCCATGCTCAC 520
```

RESULT 5
BB744515

```
LOCUS       BB744515                469 bp    mRNA    linear    EST 16-OCT-2001
DEFINITION  BB744515 RIKEN full-length enriched, adult male kidney Mus musculus
            cDNA clone F530003I24 3', mRNA sequence.
ACCESSION   BB744515
VERSION     BB744515.1  GI:16152351
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

```
REFERENCE   1  (bases 1 to 469)
  AUTHORS   Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
            Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K.,
            Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
            Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,
            Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
            2001)
  JOURNAL   Unpublished (2001)
  COMMENT   Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
            Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
            and Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
            Sugahara,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
            e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .469
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="F530003I24"
                     /sex="male"
                     /tissue_type="kidney"
                     /dev_stage="adult"
                     /lab_host="SOLR"
                     /clone_lib="RIKEN full-length enriched, adult male kidney"
                     /note="Site_1: XhoI; Site_2: SstI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
```

RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGCGGCCGCAACTCGAGTTTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGAAGGATCCAAGAGCTCAATTAATTAATTAAACCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. "

ORIGIN

```
     Query Match              29.5%;  Score 455;  DB 10;  Length 469;
     Best Local Similarity   100.0%;  Pred. No. 8.5e-99;
     Matches  455;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;


Qy      1089 TAGGCTTTAGCTTTCCACCATCCTCCAACCCCCAGGGCTGGAGTACAAGCTGGGTCCACA 1148
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 TAGGCTTTAGCTTTCCACCATCCTCCAACCCCCAGGGCTGGAGTACAAGCTGGGTCCACA 60

Qy      1149 TGAATCAGAAGGCAGCTCTCTGTTCTGATTTTAGGTTATACCCAGAGTATGGAAAAAATA 1208
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 TGAATCAGAAGGCAGCTCTCTGTTCTGATTTTAGGTTATACCCAGAGTATGGAAAAAATA 120

Qy      1209 AGGCATGAGAAAGCATTGACATCTTCACTTAAGAACTGAACAAAAGAGAACAAATATTGT 1268
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 AGGCATGAGAAAGCATTGACATCTTCACTTAAGAACTGAACAAAAGAGAACAAATATTGT 180

Qy      1269 CAATGTTTGGACACTTAGGATCTGAAATCTTGGAAATTTTAAGACCTCTTTTTCTATCAG 1328
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 CAATGTTTGGACACTTAGGATCTGAAATCTTGGAAATTTTAAGACCTCTTTTTCTATCAG 240

Qy      1329 TGTAAAAGGAATACAAGATAGCTAGTTGCAAATGCTGAATGCATTTCATCATTGGTCAGG 1388
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 TGTAAAAGGAATACAAGATAGCTAGTTGCAAATGCTGAATGCATTTCATCATTGGTCAGG 300

Qy      1389 TCGATAAGCGTGTTTCTGAAATAGTCTTATTTTTATTCTTGTAATATTAAAATTTATGTG 1448
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 TCGATAAGCGTGTTTCTGAAATAGTCTTATTTTTATTCTTGTAATATTAAAATTTATGTG 360

Qy      1449 AAAAATGAATATAATTCAATGTACAACATTAGATTTTCTATTTGAAAATTATATTTCTTG 1508
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 AAAAATGAATATAATTCAATGTACAACATTAGATTTTCTATTTGAAAATTATATTTCTTG 420

Qy      1509 AAAAAATAACTGCTGTGCCTAAATAAATCAATATA 1543
             |||||||||||||||||||||||||||||||||||
Db       421 AAAAAATAACTGCTGTGCCTAAATAAATCAATATA 455
```

RESULT 6
BB746222

```
SOURCE          Mus musculus (house mouse)
  ORGANISM      Mus musculus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE       1  (bases 1 to 458)
  AUTHORS       Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
                Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K.,
                Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
                Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,
                Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,
                Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
                Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
                Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A.,
                Muramatsu,M. and Hayashizaki,Y.
    TITLE       RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
                2001)
  JOURNAL       Unpublished (2001)
COMMENT         Contact: Yoshihide Hayashizaki
                Laboratory for Genome Exploration Research Group, RIKEN Genomic
                Sciences Center(GSC), Yokohama Institute
                The Institute of Physical and Chemical Research (RIKEN)
                1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
                Tel: 81-45-503-9222
                Fax: 81-45-503-9216
                Email: genome-res@gsc.riken.go.jp,
                URL:http://genome.gsc.riken.go.jp/
                Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
                Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
                 Normalization and subtraction of cap-trapper-selected cDNAs to
                prepare full-length cDNA libraries for rapid discovery of new
                genes. Genome Res. . 10 (10), 1617-1630 (2000)
                 wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
                Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
                Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
                and Hayashizaki,Y.
                 RIKEN integrated sequence analysis (RISA) system--384-format
                sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
                10 (11), 1757-1771 (2000)
                 Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
                Sugahara,Y. and Hayashizaki,Y.
                 Computer-based methods for the mouse full-length cDNA
                encyclopedia: real-time sequence clustering for construction of a
                nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
                 Please visit our web site (http://genome.gsc.riken.go.jp) for
                further details.
                 e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .458
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="F530013P03"
                     /sex="male"
                     /tissue_type="kidney"
                     /dev_stage="adult"
                     /lab_host="SOLR"
                     /clone_lib="RIKEN full-length enriched, adult male kidney"
```

/note="Site_1: XhoI; Site_2: SstI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGCGGCCGCAACTCGAGTTTTTTTTTTTTTTTTVN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGAGAAGGATCCAAGAGCTCAATTAATTAATTAAACCCCCCCCCCC 3'].
cDNA was cleaved with XhoI and SstI. "
ORIGIN

```
  Query Match              28.4%;   Score 438;  DB 10;   Length 458;
  Best Local Similarity    99.8%;   Pred. No. 1.1e-94;
    Matches  449;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy      1058 GTTGAGTTTTAACTAAGTAAACCACCATTTCTAGGCTTTAGCTTTCCACCATCCTCCAAC 1117
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        10 GTTGAGTTTTAACTAAGTAAACCACCATTTCTAGGCTTTAGCTTTCCACCATCCTCCAAC 69

Qy      1118 CCCCAGGGCTGGAGTACAAGCTGGGTCCACATGAATCAGAAGGCAGCTCTCTGTTCTGAT 1177
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        70 CCCCAGGGCTGGAGTACAAGCTGGGTCCACATGAATCAGAAGGCAGCTCTCTGTTCTGAT 129

Qy      1178 TTTAGGTTATACCCAGAGTATGGAAAAAATAAGGCATGAGAAAGCATTGACATCTTCACT 1237
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       130 TTTAGGTTATACCCAGAGTATGGAAAAAATAAGGCATGAGAAAGCATTGACATCTTCACT 189

Qy      1238 TAAGAACTGAACAAAAGAGAACAAATATTGTCAATGTTTGGACACTTAGGATCTGAAATC 1297
             |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       190 TAAG-ACTGAACAAAAGAGAACAAATATTGTCAATGTTTGGACACTTAGGATCTGAAATC 248

Qy      1298 TTGGAAATTTTAAGACCTCTTTTTCTATCAGTGTAAAAGGAATACAAGATAGCTAGTTGC 1357
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       249 TTGGAAATTTTAAGACCTCTTTTTCTATCAGTGTAAAAGGAATACAAGATAGCTAGTTGC 308

Qy      1358 AAATGCTGAATGCATTTCATCATTGGTCAGGTCGATAAGCGTGTTTCTGAAATAGTCTTA 1417
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       309 AAATGCTGAATGCATTTCATCATTGGTCAGGTCGATAAGCGTGTTTCTGAAATAGTCTTA 368

Qy      1418 TTTTTATTCTTGTAATATTAAAATTTATGTGAAAAATGAATATAATTCAATGTACAACAT 1477
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       369 TTTTTATTCTTGTAATATTAAAATTTATGTGAAAAATGAATATAATTCAATGTACAACAT 428

Qy      1478 TAGATTTTCTATTTGAAAATTATATTTCTT 1507
             ||||||||||||||||||||||||||||||
Db       429 TAGATTTTCTATTTGAAAATTATATTTCTT 458
```

RESULT 7
BB738743
LOCUS   BB738743   428 bp   mRNA   linear   EST 15-OCT-2001
DEFINITION   BB738743 RIKEN full-length enriched, 6 days neonate spleen Mus

musculus cDNA clone F430109C18 3', mRNA sequence.
ACCESSION   BB738743
VERSION   BB738743.1 GI:16141748
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 428)
  AUTHORS   Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
            Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K.,
            Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
            Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,
            Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE   RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
            2001)
  JOURNAL   Unpublished (2001)
  COMMENT   Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
            Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
            and Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
            Sugahara,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES   Location/Qualifiers
     source   1. .428
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="F430109C18"

/tissue_type="spleen"
            /dev_stage="6 days neonate"
            /clone_lib="RIKEN full-length enriched, 6 days neonate
            spleen"
ORIGIN

    Query Match   26.8%;   Score 414;   DB 10;   Length 428;
    Best Local Similarity   100.0%;   Pred. No. 6.5e-89;
    Matches 414; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1130 AGTACAAGCTGGGTCCACATGAATCAGAAGGCAGCTCTCTGTTCTGATTTTAGGTTATAC 1189
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 AGTACAAGCTGGGTCCACATGAATCAGAAGGCAGCTCTCTGTTCTGATTTTAGGTTATAC 60

Qy   1190 CCAGAGTATGGAAAAAATAAGGCATGAGAAAGCATTGACATCTTCACTTAAGAACTGAAC 1249
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CCAGAGTATGGAAAAAATAAGGCATGAGAAAGCATTGACATCTTCACTTAAGAACTGAAC 120

Qy   1250 AAAAGAGAACAAATATTGTCAATGTTTGGACACTTAGGATCTGAAATCTTGGAAATTTTA 1309
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AAAAGAGAACAAATATTGTCAATGTTTGGACACTTAGGATCTGAAATCTTGGAAATTTTA 180

Qy   1310 AGACCTCTTTTTCTATCAGTGTAAAAGGAATACAAGATAGCTAGTTGCAAATGCTGAATG 1369
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AGACCTCTTTTTCTATCAGTGTAAAAGGAATACAAGATAGCTAGTTGCAAATGCTGAATG 240

Qy   1370 CATTTCATCATTGGTCAGGTCGATAAGCGTGTTTCTGAAATAGTCTTATTTTTATTCTTG 1429
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CATTTCATCATTGGTCAGGTCGATAAGCGTGTTTCTGAAATAGTCTTATTTTTATTCTTG 300

Qy   1430 TAATATTAAAATTTATGTGAAAAATGAATATAATTCAATGTACAACATTAGATTTTCTAT 1489
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TAATATTAAAATTTATGTGAAAAATGAATATAATTCAATGTACAACATTAGATTTTCTAT 360

Qy   1490 TTGAAAATTATATTTCTTGAAAAAATAACTGCTGTGCCTAAATAAATCAATATA 1543
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TTGAAAATTATATTTCTTGAAAAAATAACTGCTGTGCCTAAATAAATCAATATA 414

RESULT 8
BB847918
LOCUS   BB847918   422 bp   mRNA   linear   EST 26-NOV-2001
DEFINITION   BB847918 RIKEN full-length enriched, adult male kidney Mus musculus
                cDNA clone F530201F11 5', mRNA sequence.
ACCESSION   BB847918
VERSION   BB847918.1 GI:17086293
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 422)
  AUTHORS   Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
                Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K.,
                Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
                Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,
                Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,

Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
  TITLE   RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
            2001)
  JOURNAL   Unpublished (2001)
  COMMENT   Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
            Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
            and Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
            Sugahara,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES   Location/Qualifiers
     source   1. .422
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="F530201F11"
                     /sex="male"
                     /tissue_type="kidney"
                     /dev_stage="adult"
                     /lab_host="SOLR"
                     /clone_lib="RIKEN full-length enriched, adult male kidney"
                     /note="Site_1: XhoI; Site_2: SstI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAGCGGCCGCAACTCGAGTTTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by

cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGAGAAGGATCCAAGAGCTCAATTAATTAATTAAACCCCCCCCCCC 3'].
cDNA was cleaved with XhoI and SstI. "
ORIGIN

```
   Query Match              26.2%;   Score 403.8;   DB 10;   Length 422;
   Best Local Similarity    99.5%;   Pred. No. 1.9e-86;
   Matches   405;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        1 GCTCCTGGCAGAGTTTTCTGTCGAGACAGAAGCCGACAGCAGAATGGCACAGAATTTATC 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       16 GCTCCTGGCAGAGTTTTCTGTCGAGACAGAAGCCGACAGCAGAATGGCACAGAATTTATC 75

Qy       61 TTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTA 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       76 TTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTA 135

Qy      121 TGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGTGTTCGGCTACCTCTT 180
            |||||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||
Db      136 TGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGTGTTTGGCTACCTCTT 195

Qy      181 CTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCTTTCCATCTCTGACTT 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      196 CTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCTTTCCATCTCTGACTT 255

Qy      241 TGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAATGATAAGGGGACCTA 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      256 TGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAATGATAAGGGGACCTA 315

Qy      301 TGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAACCTCTACACCAGCAT 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      316 TGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAACCTCTACACCAGCAT 375

Qy      361 CCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGCTCATGAAGT 407
            ||||||||||||||||||||||||||||||||||||||| |||||||||
Db      376 CCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGGTCATGAAGT 422
```

RESULT 9
BB864882
LOCUS   BB864882   420 bp   mRNA   linear   EST 09-JUL-2003
DEFINITION BB864882 RIKEN full-length enriched, RCB-1283 B16 melanoma cDNA Mus
            musculus cDNA clone G430047C11 5', mRNA sequence.
ACCESSION   BB864882
VERSION   BB864882.1 GI:17111092
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 420)
  AUTHORS   Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
            Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K.,
            Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
            Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,

Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.

  TITLE   RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
              2001)
  JOURNAL   Unpublished (2001)
  COMMENT   Contact: Yoshihide Hayashizaki
              Laboratory for Genome Exploration Research Group, RIKEN Genomic
              Sciences Center(GSC), Yokohama Institute
              The Institute of Physical and Chemical Research (RIKEN)
              1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
              Tel: 81-45-503-9222
              Fax: 81-45-503-9216
              Email: genome-res@gsc.riken.go.jp,
              URL:http://genome.gsc.riken.go.jp/
              Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
              Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
               Normalization and subtraction of cap-trapper-selected cDNAs to
              prepare full-length cDNA libraries for rapid discovery of new
              genes. Genome Res. . 10 (10), 1617-1630 (2000)
               wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
              Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
              Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
              and Hayashizaki,Y.
               RIKEN integrated sequence analysis (RISA) system--384-format
              sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
              10 (11), 1757-1771 (2000)
               Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
              Sugahara,Y. and Hayashizaki,Y.
               Computer-based methods for the mouse full-length cDNA
              encyclopedia: real-time sequence clustering for construction of a
              nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
               Please visit our web site (http://genome.gsc.riken.go.jp) for
              further details.
               e mouse tissues.
  FEATURES   Location/Qualifiers
       source   1. .420
                       /organism="Mus musculus"
                       /mol_type="mRNA"
                       /strain="C57BL/6"
                       /db_xref="taxon:10090"
                       /clone="G430047C11"
                       /tissue_type="skin"
                       /cell_line="RCB-1283 B16 melanoma"
                       /clone_lib="RIKEN full-length enriched, RCB-1283 B16
                       melanoma cDNA"
  ORIGIN

   Query Match   25.2%;   Score 388.4; DB 10; Length 420;
   Best Local Similarity   99.5%;   Pred. No. 9.7e-83;
   Matches 400; Conservative   0; Mismatches   1; Indels   1; Gaps   1;

Qy   1 GCTCCTGGCAGAGTTTTCTGTCGAGACAGAAGCCGACAGCAGAATGGCACAGAATTTATC 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   19 GCTCCTGGCAGAGTTTTCTGTCGAGACAGAAGCCGACAGCAGAATGGCACAGAATTTATC 78

```
Qy        61 TTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTA 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        79 TTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTA 138

Qy       121 TGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGTGTTCGGCTACCTCTT 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       139 TGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGTGTTTGGCTACCTCTT 198

Qy       181 CTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCTTTCCATCTCTGACTT 240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       199 CTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCTTTCCATCTCTGACTT 258

Qy       241 TGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAATGATAAGGGGACCTA 300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       259 TGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAATGATAAGGGGACCTA 318

Qy       301 TGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAACCTCTACACCAGCAT 360
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       319 TGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAACCTCTACACCAGCAT 378

Qy       361 CCTCTTCCTCACTTTCATTAG-CATGGACCGATATCTGCTCA 401
             |||||||||||||||||||| ||||||||||||||||||||||
Db       379 CCTCTTCCTCACTTTCATTAGCCATGGACCGATATCTGCTCA 420
```

RESULT 10
BB778587
LOCUS   BB778587   426 bp   mRNA   linear   EST 08-JUL-2003
DEFINITION   BB778587 RIKEN full-length enriched, RCB-1283 B16 melanoma cDNA Mus
                 musculus cDNA clone G430047C11 3', mRNA sequence.
ACCESSION   BB778587
VERSION   BB778587.1 GI:16939287
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus
                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 426)
  AUTHORS   Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
                 Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K.,
                 Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
                 Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,
                 Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,
                 Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
                 Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
                 Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A.,
                 Muramatsu,M. and Hayashizaki,Y.
  TITLE   RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
                 2001)
  JOURNAL   Unpublished (2001)
COMMENT   Contact: Yoshihide Hayashizaki
                 Laboratory for Genome Exploration Research Group, RIKEN Genomic
                 Sciences Center(GSC), Yokohama Institute
                 The Institute of Physical and Chemical Research (RIKEN)
                 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
                 Tel: 81-45-503-9222
                 Fax: 81-45-503-9216

  e mouse tissues.
FEATURES   Location/Qualifiers
     source   1. .426
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="G430047C11"
                     /tissue_type="skin"
                     /cell_line="RCB-1283 B16 melanoma"
                     /clone_lib="RIKEN full-length enriched, RCB-1283 B16
                     melanoma cDNA"
ORIGIN

   Query Match   24.9%;   Score 384.8; DB 10; Length 426;
   Best Local Similarity 98.8%;   Pred. No. 7.2e-82;
   Matches 419; Conservative   0; Mismatches   2; Indels   3; Gaps   3;

Qy   1123 GGGCTGGAGTACAAGCTGGGTCCACATGAATCAGAAGGCAGCTCTCTGTTCTGATTTTAG 1182
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3 GGGCTGGAGTACAAGCTGGGTCCACATGAATCAGAAGGCAGCTCTCTGTTCTGATTTTAG 62

Qy   1183 GTTATACCCAGAGTATGGAAAAAATAA-GGCATGAGAAAGCATTGACATCTTCACTTAAG 1241
             ||||||||||||||||||||||||||| ||||||| |||||||||||||||||||||||
Db   63 GTTATACCCAGAGTATGGAAAAAATAAGGGCATGAAAAAGCATTGACATCTTCACTTAAG 122

Qy   1242 AACTGAACAAAAGAGAACAAATATTGTCAATGTTTGGACACTTAGGATCTGAAATCTTGG 1301
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   123 AACTGAACAAAAGAGAACAAATATTGTCAATGTTTGGACACTTAGGATCTGAAATCTTTG 182

Qy   1302 AAATTTTAAGACCTCTTTTTCTATCAGTGTAAAAGGAATACAAGATAGCTAGTTGCAAAT 1361
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   183 AAATTTTAAGACCTCTTTTTCTATCAGTGTAAAAGGAATACAAGATAGCTAGTTGCAAAT 242

Qy   1362 GCTGAATGCATTTCATCATTGGTCA-GGTCGATAAGCGTGTTTCTGAAATAGTCTTATTT 1420
             ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||

```
Db       243  GCTGAATGCATTTCATCATTGGTCACGGTCGATAAGCGTGTTTCTGAAATAGTCTTATTT 302

Qy      1421  TTATTCTTGTAATATTAAAATTTATGTGAAAAATGAATATAATTCAATGTACAACATTAG 1480
             | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       303  TTATTCTTGTAATATTAAAATTTATGTGAAAAATGAATATAATTCAATGTACAACATTAG 362

Qy      1481  ATTTTCTA-TTTGAAAATTATATTTCTTGAAAAAATAACTGCTGTGCCTAAATAAATCAA 1539
             | | | | | | | |   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       363  ATTTTCTAGTTTGAAAATTATATTTCTTGAAAAAATAACTGCTGTGCCTAAATAAATCAA 422

Qy      1540  TATA 1543
             | | | |
Db       423  TATA 426


RESULT 11
BB739482
LOCUS       BB739482                396 bp     mRNA     linear    EST 15-OCT-2001
DEFINITION  BB739482 RIKEN full-length enriched, 6 days neonate spleen Mus
            musculus cDNA clone F430113M16 3', mRNA sequence.
ACCESSION   BB739482
VERSION     BB739482.1  GI:16142487
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 396)
  AUTHORS   Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
            Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K.,
            Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
            Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,
            Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
            2001)
  JOURNAL   Unpublished (2001)
  COMMENT   Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
            Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
```

and Hayashizaki,Y.
  RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
10 (11), 1757-1771 (2000)
  Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
  Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
  Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
         e mouse tissues.
FEATURES   Location/Qualifiers
     source   1. .396
                       /organism="Mus musculus"
                       /mol_type="mRNA"
                       /strain="C57BL/6J"
                       /db_xref="taxon:10090"
                       /clone="F430113M16"
                       /tissue_type="spleen"
                       /dev_stage="6 days neonate"
                       /clone_lib="RIKEN full-length enriched, 6 days neonate
                       spleen"
ORIGIN

  Query Match   24.7%;   Score 380.4; DB 10; Length 396;
  Best Local Similarity   99.7%;   Pred. No. 8.2e-81;
  Matches 381; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   1162 AGCTCTCTGTTCTGATTTTAGGTTATACCCAGAGTATGGAAAAAATAAGGCATGAGAAAG 1221
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 AGCTCTCTGTTCTGATTTTAGGTTATACCCAGAGTATGGAAAAAATAAGGCATGAGAAAG 60

Qy   1222 CATTGACATCTTCACTTAAGAACTGAACAAAAGAGAACAAATATTGTCAATGTTTGGACA 1281
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CATTGACATCTTCACTTAAGAACTGAACAAAAGAGAACAAATATTGTCAATGTTTGGACA 120

Qy   1282 CTTAGGATCTGAAATCTTGGAAATTTTAAGACCTCTTTTTCTATCAGTGTAAAAGGAATA 1341
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CTTAGGATCTGAAATCTTGGAAATTTTAAGACCTCTTTTTCTATCAGTGTAAAAGGAATA 180

Qy   1342 CAAGATAGCTAGTTGCAAATGCTGAATGCATTTCATCATTGGTCAGGTCGATAAGCGTGT 1401
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CAAGATAGCTAGTTGCAAATGCTGAATGCATTTCATCATTGGTCAGGTCGATAAGCGTGT 240

Qy   1402 TTCTGAAATAGTCTTATTTTTATTCTTGTAATATTAAAATTTATGTGAAAAATGAATATA 1461
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TTCTGAAATAGTCTTATTTTTATTCTTGTAATATTAAAATTTATGTGAAAAATGAATATA 300

Qy   1462 ATTCAATGTACAACATTAGATTTTCTATTTGAAAATTATATTTCTTGAAAAAATAACTGC 1521
             |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ATTCAATTTACAACATTAGATTTTCTATTTGAAAATTATATTTCTTGAAAAAATAACTGC 360

Qy   1522 TGTGCCTAAATAAATCAATATA 1543
             ||||||||||||||||||||||
Db   361 TGTGCCTAAATAAATCAATATA 382

```
RESULT 12
AI649254/c
LOCUS           AI649254                 367 bp    mRNA    linear   EST 30-APR-1999
DEFINITION      uk27c10.x1 Sugano mouse kidney mkia Mus musculus cDNA clone
                IMAGE:1970226 3', mRNA sequence.
ACCESSION       AI649254
VERSION         AI649254.1  GI:4730088
KEYWORDS        EST.
SOURCE          Mus musculus (house mouse)
  ORGANISM      Mus musculus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE       1  (bases 1 to 367)
  AUTHORS       Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
                Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
                Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
                Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
                Waterston,R. and Wilson,R.
  TITLE         The WashU-NCI Mouse EST Project 1999
  JOURNAL       Unpublished (1999)
COMMENT         Other_ESTs: uk27c10.y1
                Contact: Marra M/WashU-NCI Mouse EST Project 1999
                Washington University School of Medicine
                4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
                Tel: 314 286 1800
                Fax: 314 286 1810
                Email: mouseest@watson.wustl.edu
                This clone is available royalty-free through LLNL ; contact the
                IMAGE Consortium (info@image.llnl.gov) for further information.
                MGI:986966
                This clone was previously sequenced on the 5' end only, this new
                data is from the 3' end
                Seq primer: custom primer used
                High quality sequence stop: 353.
FEATURES             Location/Qualifiers
     source          1. .367
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1970226"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="Sugano mouse kidney mkia"
                     /note="Organ: kidney; Vector: pME18S-FL3; Site_1: DraIII
                     (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCCTTTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a  DraIII adaptor [TGTTGGCCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
                     be used to isolate the cDNA insert.  Size selection was
                     performed to exclude fragments <1.5kb.  Library
                     constructed by Dr. Sumio Sugano (University of Tokyo
                     Institute of Medical Science).  Custom primers  for
```

ORIGIN

```
  Query Match              23.6%;   Score 363.8;   DB 9;   Length 367;
  Best Local Similarity    99.5%;   Pred. No. 8.2e-77;
  Matches  365;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;
```

```
Qy     1035 CACTTGATAAACAGTGCTGTGCAGTTGAGTTTTAACTAAGTAAACCACCATTTCTAGGCT 1094
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db      367 CACTTGATAAACAGTGCTGTGCAGTTGAGTTTTAACTAAGTAAACCACCATTTCTACGCT 308

Qy     1095 TTAGCTTTCCACCATCCTCCAACCCCCAGGGCTGGAGTACAAGCTGGGTCCACATGAATC 1154
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      307 TTAGCTTTCCACCATCCTCCAACCCCCAGGGCTGGAGTACAAGCTGGGTCCACATGAATC 248

Qy     1155 AGAAGGCAGCTCTCTGTTCTGATTTTAGGTTATACCCAGAGTATGGAAAAAATAAGGCAT 1214
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      247 AGAAGGCAGCTCTCTGTTCTGATTTTAGGTTATACCCAGAGTATGGAAAAAATAAGGCAT 188

Qy     1215 GAGAAAGCATTGACATCTTCACTTAAGAACTGAACAAAAGAGAACAAATATTGTCAATGT 1274
            ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db      187 GAGAAAGCATTGACATCTTCACTTAAGATCTGAACAAAAGAGAACAAATATTGTCAATGT 128

Qy     1275 TTGGACACTTAGGATCTGAAATCTTGGAAATTTTAAGACCTCTTTTTCTATCAGTGTAAA 1334
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      127 TTGGACACTTAGGATCTGAAATCTTGGAAATTTTAAGACCTCTTTTTCTATCAGTGTAAA 68

Qy     1335 AGGAATACAAGATAGCTAGTTGCAAATGCTGAATGCATTTCATCATTGGTCAGGTCGATA 1394
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       67 AGGAATACAAGATAGCTAGTTGCAAATGCTGAATGCATTTCATCATTGGTCAGGTCGATA 8

Qy     1395 AGCGTGT 1401
            |||||||
Db        7 AGCGTGT 1
```

RESULT 13
BB645274
```
LOCUS           BB645274                636 bp    mRNA     linear   EST 31-AUG-2001
DEFINITION      BB645274 RIKEN full-length enriched, 4 days neonate male adipose
                Mus musculus cDNA clone B430012O21 5', mRNA sequence.
ACCESSION       BB645274
VERSION         BB645274.1  GI:15402306
KEYWORDS        EST.
SOURCE          Mus musculus (house mouse)
  ORGANISM      Mus musculus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE       1  (bases 1 to 636)
  AUTHORS       Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
                Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
                Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
                Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
                Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
                Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
                Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
```

TITLE   RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
JOURNAL   Unpublished (2001)
COMMENT   Contact: Yoshihide Hayashizaki
           Laboratory for Genome Exploration Research Group, RIKEN Genomic
           Sciences Center(GSC), Yokohama Institute
           The Institute of Physical and Chemical Research (RIKEN)
           1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
           Tel: 81-45-503-9222
           Fax: 81-45-503-9216
           Email: genome-res@gsc.riken.go.jp,
           URL:http://genome.gsc.riken.go.jp/
           Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
           Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Normalization and subtraction of cap-trapper-selected cDNAs to
           prepare full-length cDNA libraries for rapid discovery of new
           genes. Genome Res. . 10 (10), 1617-1630 (2000)
            wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
           Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
           Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
           and Hayashizaki,Y.
            RIKEN integrated sequence analysis (RISA) system--384-format
           sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
           10 (11), 1757-1771 (2000)
            Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
           Sugahara,Y. and Hayashizaki,Y.
            Computer-based methods for the mouse full-length cDNA
           encyclopedia: real-time sequence clustering for construction of a
           nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
            Yamanaka,I., Kiyosawa,H., Kondo,S., Saito,T., Shinagawa,A.,
           Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K.,
           Arakawa,T., Ishii,Y. and Hayashizaki,Y.
            Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct.
           Func. Genomics 2 pre, L72-L86 (2001
            Please visit our web site (http://genome.gsc.riken.go.jp) for
           further details.
            e mouse tissues.
FEATURES   Location/Qualifiers
     source   1. .636
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="B430012O21"
                     /sex="male"
                     /tissue_type="adipose"
                     /dev_stage="4 days neonate"
                     /lab_host="DH10B"
                     /clone_lib="RIKEN full-length enriched, 4 days neonate
                     male adipose"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse

transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 10.0 and subtraction to Rot = 229.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I."

ORIGIN

```
  Query Match              23.2%;  Score 357.6;  DB 10;  Length 636;
  Best Local Similarity    91.7%;  Pred. No. 2.7e-75;
   Matches 389;  Conservative   0;  Mismatches   34;  Indels    1;  Gaps    1;


Qy        1 GCTCCTGGCAGAGTTTTCTGTCGAGACAGAAGCCGACAGCAGAATGGCACAGAATTTATC 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       20 GCTCCTGGCAGAGTTTTCTGTCGAGACAGAAGCCGACAGCAGAATGGCACAGAATTTATC 79


Qy       61 TTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db       80 TTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGGATTTTA 139


Qy      121 TGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGTGTTCGGCTACCTCTT 180
            |||||||||||||||||||||||||||||||||||||||||| |||||| || ||||| ||
Db      140 TGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGGGGTGTTTGGGTACCTGTT 199


Qy      181 CTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCTTTCCATCTCTGACTT 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      200 CTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCTTTCCATCTCTGACTT 259


Qy      241 TGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAATGATAAGGGGACCTA 300
            ||||||||||| |||||||||||||||||||||||||| |||||||||||||||||||||
Db      260 TGCTTTCCTGGGCACCCTTCCCATCCTGATAAAGAGTTTTGCCAATGATAAGGGGACCTA 319


Qy      301 TGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAACCTCTACACC-AGCA 359
            |||||||||||| || |||||||||||||| |||| |||||||||| || || |||
Db      320 TGGAGATGTTCTTTGGATAAGCAACCGATATGGGCTTAACACCAACCTTTAATCCAAGCT 379


Qy      360 TCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGCTCATGAAGTACCCTTTCCGAG 419
            | || || || ||||| |||||||||||||||||| || | ||||| || || | |
Db      380 TACTTTTACTTACTTTATTTAGCATGGACCGATATTTGTTTATGAAAGTGCCCTTTTTCG 439


Qy      420 AACA 423
            || |
Db      440 AAAA 443
```


RESULT 14
BB846608
LOCUS   BB846608   416 bp   mRNA   linear   EST 26-NOV-2001
DEFINITION   BB846608 RIKEN full-length enriched, adult male kidney Mus musculus
             cDNA clone F530003I24 5', mRNA sequence.
ACCESSION   BB846608
VERSION   BB846608.1 GI:17084983
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus

```
                   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1   (bases 1 to 416)
  AUTHORS    Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
             Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K.,
             Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
             Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,
             Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,
             Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
             Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
             Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A.,
             Muramatsu,M. and Hayashizaki,Y.
    TITLE    RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
             2001)
  JOURNAL    Unpublished (2001)
  COMMENT    Contact: Yoshihide Hayashizaki
             Laboratory for Genome Exploration Research Group, RIKEN Genomic
             Sciences Center(GSC), Yokohama Institute
             The Institute of Physical and Chemical Research (RIKEN)
             1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
             Tel: 81-45-503-9222
             Fax: 81-45-503-9216
             Email: genome-res@gsc.riken.go.jp,
             URL:http://genome.gsc.riken.go.jp/
             Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
             Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
              Normalization and subtraction of cap-trapper-selected cDNAs to
             prepare full-length cDNA libraries for rapid discovery of new
             genes. Genome Res. . 10 (10), 1617-1630 (2000)
              wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
             Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
             Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
             and Hayashizaki,Y.
              RIKEN integrated sequence analysis (RISA) system--384-format
             sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
             10 (11), 1757-1771 (2000)
              Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
             Sugahara,Y. and Hayashizaki,Y.
              Computer-based methods for the mouse full-length cDNA
             encyclopedia: real-time sequence clustering for construction of a
             nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
              Please visit our web site (http://genome.gsc.riken.go.jp) for
             further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
   source            1. .416
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="F530003I24"
                     /sex="male"
                     /tissue_type="kidney"
                     /dev_stage="adult"
                     /lab_host="SOLR"
                     /clone_lib="RIKEN full-length enriched, adult male kidney"
                     /note="Site_1: XhoI; Site_2: SstI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
```

Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGCGGCCGCAACTCGAGTTTTTTTTTTTTTTTTVN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGAGAAGGATCCAAGAGCTCAATTAATTAATTAAACCCCCCCCCC 3'].
cDNA was cleaved with XhoI and SstI. "
ORIGIN

```
  Query Match              23.0%;  Score 354.2;  DB 10;  Length 416;
  Best Local Similarity    97.3%;  Pred. No. 1.7e-74;
    Matches  392;  Conservative    0;  Mismatches    8;  Indels    3;  Gaps    3;

Qy       1 GCTCCTGGCAGAGTTTTCTGTCGAGACAGAAGCCGACAGCAGAATGGCACAGAATTTATC 60
           |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db      16 GCTCCTGGCAGAGTTTTCTGTCGAGACAGAAGCCGAAAGCAGAATGGCACAGAATTTATC 75

Qy      61 TTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      76 TTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTA 135

Qy     121 TGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGTGTTCGGCTACCTCTT 180
           |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db     136 TGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGTGTTTGGCTACCTCTT 195

Qy     181 CTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCTTTCCATCTCTGACTT 240
           ||||||||||||||||||||| |||||||||||||||||||| |||| |||||||||||
Db     196 CTGCATGAAGAACTGGAAAAGCAGCAATGTCTATCTTTTTAAACTTT-CATCTCTGACTT 254

Qy     241 TGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAATGATAAGGGGACCTA 300
           |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db     255 TGCTTTCCTGTGCACCCTT-CCATCCTGATAAAGAGTTATGCCAATGATAAGGGGACCTA 313

Qy     301 TGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAACCTCTACACCAGCAT 360
           |||||||||||| |||||||||||||||||||||| ||||||| ||||||||| |||||||
Db     314 TGGAGATGTTCTATGTATAAGCAACCGATATGTGGTTCACAACAACCTCTAAACCAGCAT 373

Qy     361 CCTCTTCCTCACTTTCATTAG-CATGGACCGATATCTGCTCAT 402
           |||||||||||||||||||||| |||||||||||||||||||||
Db     374 CCTCTTCCTCACTTTCATTAGCCATGGACCGATATCTGCTCAT 416
```

RESULT 15
BY368584
LOCUS   BY368584   408 bp   mRNA   linear   EST 12-DEC-2002
DEFINITION BY368584 RIKEN full-length enriched, 6 days neonate spleen Mus
            musculus cDNA clone F430110C01 3', mRNA sequence.
ACCESSION   BY368584
VERSION   BY368584.1 GI:26598072
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus

```
                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
     REFERENCE      1  (bases 1 to 408)
       AUTHORS      Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
                    Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
                    Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
                    Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
                    Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
                    Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
                    Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
                    Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
                    Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
                    Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
                    Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
                    Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
                    Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
                    Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
                    Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
                    Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
                    Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
                    Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
                    Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
                    Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
                    Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
                    Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
                    Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
                    Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
                    Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
                    Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
                    Rogers,J., Birney,E. and Hayashizaki,Y.
        TITLE       Analysis of the mouse transcriptome based on functional annotation
                    of 60,770 full-length cDNAs
      JOURNAL       Nature 420, 563-573 (2002)
      MEDLINE       22354683
       PUBMED       12466851
      COMMENT       Contact: Yoshihide Hayashizaki
                    Laboratory for Genome Exploration Research Group, RIKEN Genomic
                    Sciences Center(GSC), Yokohama Institute
                    The Institute of Physical and Chemical Research (RIKEN)
                    1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
                    Tel: 81-45-503-9222
                    Fax: 81-45-503-9216
                    Email: genome-res@gsc.riken.go.jp,
                    URL:http://genome.gsc.riken.go.jp/
                    Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,
                    Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H.,
                    Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R.,
                    Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K.,
                    Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and
                    Hayashizaki,Y. Direct Submission
                     Computational Analysis of Full-Length Mouse cDNAs Compared with
                    Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
                     Normalization and subtraction of cap-trapper-selected cDNAs to
                    prepare full-length cDNA libraries for rapid discovery of new
                    genes. Genome Res. 10 (10), 1617-1630 (2000)
                     RIKEN integrated sequence analysis (RISA) system--384-format
                    sequencing pipeline with 384 multicapillary sequencer. Genome Res.
```

10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
```
FEATURES             Location/Qualifiers
     source          1. .408
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="F430110C01"
                     /tissue_type="spleen"
                     /dev_stage="6 days neonate"
                     /clone_lib="RIKEN full-length enriched, 6 days neonate
                     spleen"
ORIGIN
```

```
  Query Match              22.7%;   Score 350.6;   DB 13;   Length 408;
  Best Local Similarity    98.0%;   Pred. No. 1.3e-73;
  Matches  386;  Conservative    0;  Mismatches    5;  Indels    3;  Gaps    3;


Qy     1153 TCAGAAGGCAGCTCTCTGTTCTGATTTTAGGTTATACCCAGAGTATGGAAAAAATAAGGC 1212
            |||||||  ||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        1 TCAGAAGCCAGCTCTCTGTTCTGATTTTAGGTTATACCCAGAGTATGNAAAAAATAAGGC 60


Qy     1213 ATGAGAAAGCA-TTGACATCTTCACTTAAGAACTGAACAAAAGAGAACAAATA-TTGTCA 1270
            ||||||||||| ||||||||||||||||||||||||||||||||||||||||| || |||
Db       61 ATGAGAAAGCAGTTGACATCTTCACTTAAGAACTGAACAAAAGAGAACAAATAGTTCTCA 120


Qy     1271 ATGTTTGGACACTTAGGATCTGAAATCTTGGAAATTTTAAGACCTCTTTTTCTATCAGTG 1330
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 ATGTTTGGACACTTAGGATCTGAAATCTTGGAAATTTTAAGACCTCTTTTTCTATCAGTG 180


Qy     1331 TAAAAGGAATACAAGATAG-CTAGTTGCAAATGCTGAATGCATTTCATCATTGGTCAGGT 1389
            ||||||||||||||||||| |||||| |||||||||||||||||||||||||||||||||
Db      181 TAAAAGGAATACAAGATAGCCTAGTTCCAAATGCTGAATGCATTTCATCATTGGTCAGGT 240


Qy     1390 CGATAAGCGTGTTTCTGAAATAGTCTTATTTTTATTCTTGTAATATTAAAATTTATGTGA 1449
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 CGATAAGCGTGTTTCTGAAATAGTCTTATTTTTATTCTTGTAATATTAAAATTTATGTGA 300


Qy     1450 AAAATGAATATAATTCAATGTACAACATTAGATTTTCTATTTGAAAATTATATTTCTTGA 1509
            ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db      301 AAAATGAATATAATTCAATTTACAACATTAGATTTTCTATTTGAAAATTATATTTCTTGA 360


Qy     1510 AAAAATAACTGCTGTGCCTAAATAAATCAATATA 1543
            ||||||||||||||||||||||||||||||||||
Db      361 AAAAATAACTGCTGTGCCTAAATAAATCAATATA 394
```

Search completed: August 24, 2004, 16:03:07
Job time : 4316 secs

```
                        GenCore version 5.1.6
                Copyright (c) 1993 - 2004  Compugen Ltd.


OM nucleic - nucleic search, using sw model

Run on:          August 24, 2004, 08:45:05 ; Search time 6253 Seconds
                                            (without alignments)
                                            10695.394 Million cell updates/sec

Title:           US-09-891-138A-1
Perfect score:   1543
Sequence:        1 gctcctggcagagttttctg..........tgcctaaataaatcaatata 1543

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters:      6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       GenEmbl:*
                 1:  gb_ba:*
                 2:  gb_htg:*
                 3:  gb_in:*
                 4:  gb_om:*
                 5:  gb_ov:*
                 6:  gb_pat:*
                 7:  gb_ph:*
                 8:  gb_pl:*
                 9:  gb_pr:*
                 10: gb_ro:*
                 11: gb_sts:*
                 12: gb_sy:*
                 13: gb_un:*
                 14: gb_vi:*
                 15: em_ba:*
                 16: em_fun:*
                 17: em_hum:*
                 18: em_in:*
                 19: em_mu:*
                 20: em_om:*
                 21: em_or:*
                 22: em_ov:*
                 23: em_pat:*
                 24: em_ph:*
                 25: em_pl:*
                 26: em_ro:*
                 27: em_sts:*
```

```
28:  em_un:*
29:  em_vi:*
30:  em_htg_hum:*
31:  em_htg_inv:*
32:  em_htg_other:*
33:  em_htg_mus:*
34:  em_htg_pln:*
35:  em_htg_rod:*
36:  em_htg_mam:*
37:  em_htg_vrt:*
38:  em_sy:*
39:  em_htgo_hum:*
40:  em_htgo_mus:*
41:  em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1543	100.0	1543	6	AX376573	AX376573 Sequence
	2	1533.4	99.4	1598	10	AF295367	AF295367 Mus muscu
c	3	1494.8	96.9	202487	2	AC138318	AC138318 Mus muscu
c	4	1041.6	67.5	239576	2	AC111231	AC111231 Rattus no
	5	645.8	41.9	60298	2	AC116149	AC116149 Mus muscu
c	6	633.6	41.1	60298	2	AC116149	AC116149 Mus muscu
c	7	615.8	39.9	326606	2	AC110839	AC110839 Rattus no
	8	592.4	38.4	1325	9	AF247785	AF247785 Homo sapi
	9	592.4	38.4	1380	6	AX549281	AX549281 Sequence
	10	592.4	38.4	1380	6	AX780453	AX780453 Sequence
	11	592.4	38.4	1380	9	AF348078	AF348078 Homo sapi
	12	592.4	38.4	1449	9	BC030948	BC030948 Homo sapi
	13	592.4	38.4	1542	6	AX342665	AX342665 Sequence
	14	590.2	38.3	90343	9	AC116026	AC116026 Homo sapi
	15	590.2	38.3	132745	9	AC068647	AC068647 Homo sapi
	16	589.2	38.2	1996	6	AR035943	AR035943 Sequence
c	17	588.6	38.1	158144	2	AC016455	AC016455 Homo sapi
	18	572.8	37.1	1005	6	AX657428	AX657428 Sequence
c	19	158.8	10.3	639	6	AX775477	AX775477 Sequence
	20	147.6	9.6	1977	5	AF031897	AF031897 Meleagris
	21	145.2	9.4	69462	2	AC101335	AC101335 Mus muscu
	22	133.8	8.7	251	6	AX779053	AX779053 Sequence
	23	126.6	8.2	1011	6	BD171271	BD171271 Novel G p
	24	126.6	8.2	1011	6	BD184762	BD184762 Novel G p
	25	126.6	8.2	1014	6	AX148186	AX148186 Sequence
	26	126.6	8.2	1014	6	AX379468	AX379468 Sequence
	27	126.6	8.2	1014	6	AX384211	AX384211 Sequence
	28	126.6	8.2	1014	6	AX549361	AX549361 Sequence
	29	126.6	8.2	1014	6	AX593341	AX593341 Sequence
	30	126.6	8.2	1014	6	BD144282	BD144282 Novel G-p
	31	126.6	8.2	1014	9	AB083598	AB083598 Homo sapi
	32	126.6	8.2	1014	9	AF411109	AF411109 Homo sapi
	33	126.6	8.2	1029	6	AX593346	AX593346 Sequence

```
    34    126.6    8.2    1081   6   AX458238              AX458238  Sequence
    35    126.6    8.2    1092   9   AF370886              AF370886  Homo sapi
    36    126.6    8.2    1179   6   AX593351              AX593351  Sequence
    37    126.6    8.2    1414   6   AX646271              AX646271  Sequence
    38    126.6    8.2    1414   9   AB065877              AB065877  Homo sapi
    39    126.6    8.2    1729   6   AX191332              AX191332  Sequence
    40    126.6    8.2    1797   6   AX593340              AX593340  Sequence
    41    126.6    8.2    9905   6   AX379470              AX379470  Sequence
 c  42    126.6    8.2   67645   9   AL356486              AL356486  Human DNA
    43    126.6    8.2  156555   9   AC026756              AC026756  Homo sapi
    44    125.6    8.1     990   6   AX657422              AX657422  Sequence
    45      125    8.1    1014   6   AX305130              AX305130  Sequence
```

ALIGNMENTS

RESULT 1
AX376573
LOCUS   AX376573   1543 bp   DNA   linear   PAT 01-MAR-2002
DEFINITION   Sequence 1 from Patent WO0200719.
ACCESSION   AX376573
VERSION   AX376573.1 GI:19170674
KEYWORDS   .
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Lin,D.C., Zhao,J., Chen,J.L. and Cutler,G.
  TITLE   Novel receptors
  JOURNAL   Patent: WO 0200719-A 1 03-JAN-2002;
             Tularik Inc. (US)
FEATURES   Location/Qualifiers
     source   1. .1543
                     /organism="Mus musculus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10090"
     CDS   44. .997
                     /note="unnamed protein product; mouse TGR18 G-protein
                     coupled receptor (GPCR)"
                     /codon_start=1
                     /protein_id="CAD26816.1"
                     /db_xref="GI:19170675"
                     /db_xref="REMTREMBL:CAD26816"
                     /translation="MAQNLSCENWLATEAILNKYYLSAFYAIEFIFGLLGNVTVVFGY
                     LFCMKNWNSSNVYLFNLSISDFAFLCTLPILIKSYANDKGTYGDVLCISNRYVLHTNL
                     YTSILFLTFISMDRYLLMKYPFREHFLQKKEFAILISLAVWALVTLEVLPMLTFINSV
                     PKEEGSNCIDYASSGNPEHNLIYSLCLTLLGFLIPLSVMCFFYYKMVVFLKRRSQQQA
                     TALPLDKPQRLVVLAVVIFSILFTPYHIMRNLRIASRLDSWPQGCTQKAIKSIYTLTR
                     PLAFLNSAINPIFYFLMGDHYREMLISKFRQYFKSLTSFRT"
ORIGIN

     Query Match   100.0%; Score 1543; DB 6; Length 1543;
     Best Local Similarity 100.0%; Pred. No. 0;
     Matches 1543; Conservative   0; Mismatches   0; Indels   0; Gaps   0;
```

```
Qy     1 GCTCCTGGCAGAGTTTTCTGTCGAGACAGAAGCCGACAGCAGAATGGCACAGAATTTATC 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 GCTCCTGGCAGAGTTTTCTGTCGAGACAGAAGCCGACAGCAGAATGGCACAGAATTTATC 60

Qy    61 TTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTA 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTA 120

Qy   121 TGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGTGTTCGGCTACCTCTT 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGTGTTCGGCTACCTCTT 180

Qy   181 CTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCTTTCCATCTCTGACTT 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCTTTCCATCTCTGACTT 240

Qy   241 TGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAATGATAAGGGGACCTA 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAATGATAAGGGGACCTA 300

Qy   301 TGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAACCTCTACACCAGCAT 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAACCTCTACACCAGCAT 360

Qy   361 CCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGCTCATGAAGTACCCTTTCCGAGA 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGCTCATGAAGTACCCTTTCCGAGA 420

Qy   421 ACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCTCGCTGGCTGTCTGGGCCTTAGT 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 ACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCTCGCTGGCTGTCTGGGCCTTAGT 480

Qy   481 GACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATTCTGTCCCAAAAGAAGAGGGCAG 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   481 GACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATTCTGTCCCAAAAGAAGAGGGCAG 540

Qy   541 TAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAACACAATCTCATTTACAGCCTCTG 600
         ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   541 TAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAACACAATCTCATTTACAGCCTCTG 600

Qy   601 CCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGATGTGCTTCTTCTACTACAAGAT 660
         |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   601 CCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGATGTGCTTCTTCTACTACAAGAT 660

Qy   661 GGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAACTGCCCTGCCACTGGACAAACC 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAACTGCCCTGCCACTGGACAAACC 720

Qy   721 CCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTATACTCTTCACACCCTATCATAT 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTATACTCTTCACACCCTATCATAT 780

Qy   781 CATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTTGGCCACAAGGATGTACACAGAA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTTGGCCACAAGGATGTACACAGAA 840

Qy   841 GGCCATCAAATCTATATACACACTGACACGGCCTCTGGCCTTTCTGAACAGTGCCATCAA 900
```

```
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841  GGCCATCAAATCTATATACACACTGACACGGCCTCTGGCCTTTCTGAACAGTGCCATCAA  900

Qy     901  TCCCATCTTCTACTTCCTCATGGGAGACCATTACAGAGAGATGCTGATTAGTAAGTTCAG  960
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901  TCCCATCTTCTACTTCCTCATGGGAGACCATTACAGAGAGATGCTGATTAGTAAGTTCAG  960

Qy     961  ACAATACTTCAAGTCCCTTACATCCTTCAGGACATGAGCTGCTGGATGCAGGTCTTCACT  1020
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961  ACAATACTTCAAGTCCCTTACATCCTTCAGGACATGAGCTGCTGGATGCAGGTCTTCACT  1020

Qy    1021  CAGCCAAAATGAGACACTTGATAAACAGTGCTGTGCAGTTGAGTTTTAACTAAGTAAACC  1080
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021  CAGCCAAAATGAGACACTTGATAAACAGTGCTGTGCAGTTGAGTTTTAACTAAGTAAACC  1080

Qy    1081  ACCATTTCTAGGCTTTAGCTTTCCACCATCCTCCAACCCCCAGGGCTGGAGTACAAGCTG  1140
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081  ACCATTTCTAGGCTTTAGCTTTCCACCATCCTCCAACCCCCAGGGCTGGAGTACAAGCTG  1140

Qy    1141  GGTCCACATGAATCAGAAGGCAGCTCTCTGTTCTGATTTTAGGTTATACCCAGAGTATGG  1200
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141  GGTCCACATGAATCAGAAGGCAGCTCTCTGTTCTGATTTTAGGTTATACCCAGAGTATGG  1200

Qy    1201  AAAAAATAAGGCATGAGAAAGCATTGACATCTTCACTTAAGAACTGAACAAAAGAGAACA  1260
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201  AAAAAATAAGGCATGAGAAAGCATTGACATCTTCACTTAAGAACTGAACAAAAGAGAACA  1260

Qy    1261  AATATTGTCAATGTTTGGACACTTAGGATCTGAAATCTTGGAAATTTTAAGACCTCTTTT  1320
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261  AATATTGTCAATGTTTGGACACTTAGGATCTGAAATCTTGGAAATTTTAAGACCTCTTTT  1320

Qy    1321  TCTATCAGTGTAAAAGGAATACAAGATAGCTAGTTGCAAATGCTGAATGCATTTCATCAT  1380
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321  TCTATCAGTGTAAAAGGAATACAAGATAGCTAGTTGCAAATGCTGAATGCATTTCATCAT  1380

Qy    1381  TGGTCAGGTCGATAAGCGTGTTTCTGAAATAGTCTTATTTTTATTCTTGTAATATTAAAA  1440
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381  TGGTCAGGTCGATAAGCGTGTTTCTGAAATAGTCTTATTTTTATTCTTGTAATATTAAAA  1440

Qy    1441  TTTATGTGAAAAATGAATATAATTCAATGTACAACATTAGATTTTCTATTTGAAAATTAT  1500
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441  TTTATGTGAAAAATGAATATAATTCAATGTACAACATTAGATTTTCTATTTGAAAATTAT  1500

Qy    1501  ATTTCTTGAAAAAATAACTGCTGTGCCTAAATAAATCAATATA  1543
                  |||||||||||||||||||||||||||||||||||||||||||
Db    1501  ATTTCTTGAAAAAATAACTGCTGTGCCTAAATAAATCAATATA  1543
```

```
RESULT 2
AF295367
LOCUS       AF295367              1598 bp    mRNA    linear    ROD 06-APR-2001
DEFINITION  Mus musculus G-protein coupled receptor GPR91 mRNA, complete cds.
ACCESSION   AF295367
VERSION     AF295367.1  GI:12711490
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
```

```
     Query Match                  99.4%;   Score 1533.4;   DB 10;   Length 1598;
     Best Local Similarity        99.6%;   Pred. No. 0;
     Matches 1537;  Conservative      0;  Mismatches      6;  Indels    0;  Gaps      0;


Qy         1 GCTCCTGGCAGAGTTTTCTGTCGAGACAGAAGCCGACAGCAGAATGGCACAGAATTTATC 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        31 GCTCCTGGCAGAGTTTTCTGTCGAGACAGAAGCCGACAGCAGAATGGCACAGAATTTATC 90

Qy        61 TTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTA 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        91 TTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTA 150

Qy       121 TGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGTGTTCGGCTACCTCTT 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       151 TGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGTGTTTGGCTACCTCTT 210

Qy       181 CTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCTTTCCATCTCTGACTT 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       211 CTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCTTTCCATCTCTGACTT 270

Qy       241 TGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAATGATAAGGGGACCTA 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       271 TGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAATGATAAGGGGACCTA 330
```

```
Qy     301 TGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAACCTCTACACCAGCAT 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     331 TGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAACCTCTACACCAGCAT 390

Qy     361 CCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGCTCATGAAGTACCCTTTCCGAGA 420
           |||||| ||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db     391 GCTCTTGCTCACTGTCATTAGCATGGACCGATATCTGCTCATGAAGTACCCTTTCCGAGA 450

Qy     421 ACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCTCGCTGGCTGTCTGGGCCTTAGT 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     451 ACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCTCGCTGGCTGTCTGGGCCTTAGT 510

Qy     481 GACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATTCTGTCCCAAAAGAAGAGGGCAG 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     511 GACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATTCTGTCCCAAAAGAAGAGGGCAG 570

Qy     541 TAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAACACAATCTCATTTACAGCCTCTG 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     571 TAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAACACAATCTCATTTACAGCCTCTG 630

Qy     601 CCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGATGTGCTTCTTCTACTACAAGAT 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     631 CCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGATGTGCTTCTTCTACTACAAGAT 690

Qy     661 GGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAACTGCCCTGCCACTGGACAAACC 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     691 GGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAACTGCCCTGCCACTGGACAAACC 750

Qy     721 CCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTATACTCTTCACACCCTATCATAT 780
           ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db     751 CCAACGCCTGGTGGTCCTGGCAGTTGTGATCTTCTCTATACTCTTCACACCCTATCATAT 810

Qy     781 CATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTTGGCCACAAGGATGTACACAGAA 840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     811 CATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTTGGCCACAAGGATGTACACAGAA 870

Qy     841 GGCCATCAAATCTATATACACACTGACACGGCCTCTGGCCTTTCTGAACAGTGCCATCAA 900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     871 GGCCATCAAATCTATATACACACTGACACGGCCTCTGGCCTTTCTGAACAGTGCCATCAA 930

Qy     901 TCCCATCTTCTACTTCCTCATGGGAGACCATTACAGAGAGATGCTGATTAGTAAGTTCAG 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     931 TCCCATCTTCTACTTCCTCATGGGAGACCATTACAGAGAGATGCTGATTAGTAAGTTCAG 990

Qy     961 ACAATACTTCAAGTCCCTTACATCCTTCAGGACATGAGCTGCTGGATGCAGGTCTTCACT 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     991 ACAATACTTCAAGTCCCTTACATCCTTCAGGACATGAGCTGCTGGATGCAGGTCTTCACT 1050

Qy    1021 CAGCCAAAATGAGACACTTGATAAACAGTGCTGTGCAGTTGAGTTTTAACTAAGTAAACC 1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1051 CAGCCAAAATGAGACACTTGATAAACAGTGCTGTGCAGTTGAGTTTTAACTAAGTAAACC 1110

Qy    1081 ACCATTTCTAGGCTTTAGCTTTCCACCATCCTCCAACCCCCAGGGCTGGAGTACAAGCTG 1140
           ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db    1111 ACCATTTCTACGCTTTAGCTTTCCACCATCCTCCAACCCCCAGGGCTGGAGTACAAGCTG 1170
```

```
Qy   1141 GGTCCACATGAATCAGAAGGCAGCTCTCTGTTCTGATTTTAGGTTATACCCAGAGTATGG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1171 GGTCCACATGAATCAGAAGGCAGCTCTCTGTTCTGATTTTAGGTTATACCCAGAGTATGG 1230

Qy   1201 AAAAAATAAGGCATGAGAAAGCATTGACATCTTCACTTAAGAACTGAACAAAAGAGAACA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1231 AAAAAATAAGGCATGAGAAAGCATTGACATCTTCACTTAAGAACTGAACAAAAGAGAACA 1290

Qy   1261 AATATTGTCAATGTTTGGACACTTAGGATCTGAAATCTTGGAAATTTTAAGACCTCTTTT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1291 AATATTGTCAATGTTTGGACACTTAGGATCTGAAATCTTGGAAATTTTAAGACCTCTTTT 1350

Qy   1321 TCTATCAGTGTAAAAGGAATACAAGATAGCTAGTTGCAAATGCTGAATGCATTTCATCAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1351 TCTATCAGTGTAAAAGGAATACAAGATAGCTAGTTGCAAATGCTGAATGCATTTCATCAT 1410

Qy   1381 TGGTCAGGTCGATAAGCGTGTTTCTGAAATAGTCTTATTTTTATTCTTGTAATATTAAAA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1411 TGGTCAGGTCGATAAGCGTGTTTCTGAAATAGTCTTATTTTTATTCTTGTAATATTAAAA 1470

Qy   1441 TTTATGTGAAAAATGAATATAATTCAATGTACAACATTAGATTTTCTATTTGAAAATTAT 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1471 TTTATGTGAAAAATGAATATAATTCAATGTACAACATTAGATTTTCTATTTGAAAATTAT 1530

Qy   1501 ATTTCTTGAAAAAATAACTGCTGTGCCTAAATAAATCAATATA 1543
          |||||||||||||||||||||||||||||||||||||||||||
Db   1531 ATTTCTTGAAAAAATAACTGCTGTGCCTAAATAAATCAATATA 1573


RESULT 3
AC138318/c
LOCUS       AC138318              202487 bp    DNA     linear   HTG 18-DEC-2003
DEFINITION  Mus musculus chromosome 3 clone RP23-358I23 map 3, *** SEQUENCING
            IN PROGRESS ***, 7 unordered pieces.
ACCESSION   AC138318
VERSION     AC138318.4  GI:40018777
KEYWORDS    HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 202487)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus chromosome 3, clone RP23-358I23
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 202487)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
            Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
```

```
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
            Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
            Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
```
  TITLE   Direct Submission
  JOURNAL   Submitted (25-DEC-2002) Whitehead Institute/MIT Center for Genome
             Research, 320 Charles Street, Cambridge, MA 02141, USA
  REFERENCE 3 (bases 1 to 202487)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
             Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
             Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
             Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
             Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
             Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
             Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
             Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
             Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
             Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C.,
             Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
             Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
             Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
             O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
             Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
             Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
             Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
             Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
             Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
             Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (18-DEC-2003) Whitehead Institute/MIT Center for Genome
             Research, 320 Charles Street, Cambridge, MA 02141, USA
  COMMENT   On Dec 18, 2003 this sequence version replaced gi:29150501.
             All repeats were identified using RepeatMasker:
             Smit, A.F.A. & Green, P. (1996-1997)
             http://ftp.genome.washington.edu/RM/RepeatMasker.html
             -------------- Genome Center
                 Center: Whitehead Institute/ MIT Center for Genome Research
                 Center code: WIBR
                 Web site: http://www-seq.wi.mit.edu
                 Contact: sequence_submissions@genome.wi.mit.edu
             -------------- Project Information
                 Center project name: L28921
                 Center clone name: 358_I_23
             --------------.
             * NOTE: This is a 'working draft' sequence. It currently
             * consists of 7 contigs. The true order of the pieces
             * is not known and their order in this sequence record is
             * arbitrary. Gaps between the contigs are represented as
             * runs of N, but the exact sizes of the gaps are unknown.
             * This record will be updated with the finished sequence
             * as soon as it is available and the accession number will
             * be preserved.
             *   1   68110: contig of 68110 bp in length
             *   68111   68210: gap of 100 bp

```
     *      68211      72920: contig of 4710 bp in length
     *      72921      73020: gap of 100 bp
     *      73021     125850: contig of 52830 bp in length
     *     125851     125950: gap of 100 bp
     *     125951     142368: contig of 16418 bp in length
     *     142369     142468: gap of 100 bp
     *     142469     146579: contig of 4111 bp in length
     *     146580     146679: gap of 100 bp
     *     146680     162672: contig of 15993 bp in length
     *     162673     162772: gap of 100 bp
     *     162773     202487: contig of 39715 bp in length.
FEATURES                Location/Qualifiers
     source             1. .202487
                        /organism="Mus musculus"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:10090"
                        /chromosome="3"
                        /map="3"
                        /clone="RP23-358I23"
                        /clone_lib="RPCI-23 Female Mouse BAC"
ORIGIN

  Query Match            96.9%;  Score 1494.8;  DB 2;  Length 202487;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches 1496;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;


Qy        46 GGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTA 105
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105379 GGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTA
      105320


Qy       106 CCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGT 165
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105319 CCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGT
      105260


Qy       166 GTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCT 225
             |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105259 GTTTGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCT
      105200


Qy       226 TTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAA 285
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105199 TTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAA
      105140


Qy       286 TGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAA 345
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105139 TGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAA
      105080


Qy       346 CCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGCTCATGAA 405
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105079 CCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGCTCATGAA
      105020


Qy       406 GTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCTCGCTGGC 465
```

```
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105019  GTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCTCGCTGGC
104960


Qy       466  TGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATTCTGTCCC 525
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    104959  TGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATTCTGTCCC
104900


Qy       526  AAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAACACAATCT 585
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    104899  AAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAACACAATCT
104840


Qy       586  CATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTGTGATGTGCTT 645
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    104839  CATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTGTGATGTGCTT
104780


Qy       646  CTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAACTGCCCT 705
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    104779  CTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAACTGCCCT
104720


Qy       706  GCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTATACTCTT 765
              ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db    104719  GCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCAGTTGTGATCTTCTCTATACTCTT
104660


Qy       766  CACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTTGGCCACA 825
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    104659  CACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTTGGCCACA
104600


Qy       826  AGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTCTGGCCTTTCT 885
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    104599  AGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTCTGGCCTTTCT
104540


Qy       886  GAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACAGAGAGATGCT 945
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    104539  GAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACAGAGAGATGCT
104480


Qy       946  GATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACATGAGCTGCTGG 1005
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    104479  GATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACATGAGCTGCTGG
104420


Qy      1006  ATGCAGGTCTTCACTCAGCCAAAATGAGACACTTGATAAACAGTGCTGTGCAGTTGAGTT 1065
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    104419  ATGCAGGTCTTCACTCAGCCAAAATGAGACACTTGATAAACAGTGCTGTGCAGTTGAGTT
104360


Qy      1066  TTAACTAAGTAAACCACCATTTCTAGGCTTTAGCTTTCCACCATCCTCCAACCCCCAGGG 1125
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    104359 TTAACTAAGTAAACCACCATTTCTAGGCTTTAGCTTTCCACCATCCTCCAACCCCCAGGG
      104300

Qy      1126 CTGGAGTACAAGCTGGGTCCACATGAATCAGAAGGCAGCTCTCTGTTCTGATTTTAGGTT 1185
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    104299 CTGGAGTACAAGCTGGGTCCACATGAATCAGAAGGCAGCTCTCTGTTCTGATTTTAGGTT
      104240

Qy      1186 ATACCCAGAGTATGGAAAAAATAAGGCATGAGAAAGCATTGACATCTTCACTTAAGAACT 1245
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    104239 ATACCCAGAGTATGGAAAAAATAAGGCATGAGAAAGCATTGACATCTTCACTTAAGAACT
      104180

Qy      1246 GAACAAAAGAGAACAAATATTGTCAATGTTTGGACACTTAGGATCTGAAATCTTGGAAAT 1305
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    104179 GAACAAAAGAGAACAAATATTGTCAATGTTTGGACACTTAGGATCTGAAATCTTGGAAAT
      104120

Qy      1306 TTTAAGACCTCTTTTTCTATCAGTGTAAAAGGAATACAAGATAGCTAGTTGCAAATGCTG 1365
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    104119 TTTAAGACCTCTTTTTCTATCAGTGTAAAAGGAATACAAGATAGCTAGTTGCAAATGCTG
      104060

Qy      1366 AATGCATTTCATCATTGGTCAGGTCGATAAGCGTGTTTCTGAAATAGTCTTATTTTTATT 1425
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    104059 AATGCATTTCATCATTGGTCAGGTCGATAAGCGTGTTTCTGAAATAGTCTTATTTTTATT
      104000

Qy      1426 CTTGTAATATTAAAATTTATGTGAAAAATGAATATAATTCAATGTACAACATTAGATTTT 1485
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    103999 CTTGTAATATTAAAATTTATGTGAAAAATGAATATAATTCAATGTACAACATTAGATTTT
      103940

Qy      1486 CTATTTGAAAATTATATTTCTTGAAAAAATAACTGCTGTGCCTAAATAAATCAATATA 1543
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    103939 CTATTTGAAAATTATATTTCTTGAAAAAATAACTGCTGTGCCTAAATAAATCAATATA 103882
```

RESULT 4
AC111231/c
LOCUS   AC111231   239576 bp   DNA   linear   HTG 13-MAY-2003
DEFINITION Rattus norvegicus clone CH230-96O13, *** SEQUENCING IN PROGRESS
            ***, 2 unordered pieces.
ACCESSION   AC111231
VERSION   AC111231.7 GI:30578486
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE   Rattus norvegicus (Norway rat)
  ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 239576)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,

Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
  TITLE   Direct Submission
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 239576)
  AUTHORS   Worley,K.C.
  TITLE   Direct Submission
  JOURNAL   Submitted (19-FEB-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3 (bases 1 to 239576)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE   Direct Submission
  JOURNAL   Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One

```
                       Baylor Plaza, Houston, TX 77030, USA
COMMENT         On May 13, 2003 this sequence version replaced gi:24819079.
                The sequence in this assembly is a combination of BAC based reads
                and whole genome shotgun sequencing reads assembled using Atlas
                (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
                in the feature table below represents a scaffold in the Atlas
                assembly (a 'contig-scaffold'). Within each contig-scaffold,
                individual sequence contigs are ordered and oriented, and separated
                by sized gaps filled with Ns to the estimated size. The sequence
                may extend beyond the ends of the clone and there may be sequence
                contigs within a contig-scaffold that consist entirely of whole
                genome shotgun sequence reads. Both end sequences and whole genome
                shotgun sequence only contigs will be indicated in the feature
                table.
                -------------- Genome Center
                    Center: Baylor College of Medicine
                    Center code: BCM
                    Web site: http://www.hgsc.bcm.tmc.edu/
                    Contact: hgsc-help@bcm.tmc.edu
                -------------- Project Information
                    Center project name: GLVO
                    Center clone name: CH230-96O13
                -------------- Summary Statistics
                    Assembly program: Atlas 3.0;
                    Consensus quality: 213738 bases at least Q40
                    Consensus quality: 217471 bases at least Q30
                    Consensus quality: 220066 bases at least Q20
                    Estimated insert size: 227472; sum-of-contigs estimation
                    Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
                --------------
                * NOTE: Estimated insert size may differ from sequence length
                *    (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
                * NOTE: This is a 'working draft' sequence. It currently
                * consists of 2 contigs. The true order of the pieces
                * is not known and their order in this sequence record is
                * arbitrary. Gaps between the contigs are represented as
                * runs of N, but the exact sizes of the gaps are unknown.
                * This record will be updated with the finished sequence
                * as soon as it is available and the accession number will
                * be preserved.
                *        1    236521: contig of 236521 bp in length
                *   236522    236621: gap of unknown length
                *   236622    239576: contig of 2955 bp in length.
FEATURES             Location/Qualifiers
     source          1. .239576
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-96O13"
     misc_feature    157219. .158900
                     /note="wgs_contig"
     misc_feature    206334. .207349
                     /note="wgs_contig"
ORIGIN

   Query Match              67.5%;  Score 1041.6;  DB 2;  Length 239576;
   Best Local Similarity    85.2%;  Pred. No. 1.2e-221;
```

Matches 1287; Conservative   0; Mismatches 194; Indels 29; Gaps 10;

```
Qy       46 GGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTA 105
            |||||||||||||||||||||||||| |||||| ||||| |||| ||| |||||| ||||||||
Db    92574 GGCACAGAATTTATCTTGTGAAAATTGGCTGGCATTAGAGAATATTTTGAAAAGTACTA 92515

Qy      106 CCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGT 165
            ||||||||||||||||  |||||||||||| ||||| ||||||| ||| |||| |||||
Db    92514 CCTCTCTGCATTTTATGGGATCGAGTTCATTGTTGGAATGCTTGGCAATTTCACCGTGGT 92455

Qy      166 GTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCT 225
            |||||||||||||||||||||||||||||||||||||||| ||||| ||||||||| || |||||
Db    92454 GTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGTAGCAACGTCTATCTCTTCAACCT 92395

Qy      226 TTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAA 285
            |||||||||||||| |||||||||||||||| |||||||| ||||||| ||||||| ||||
Db    92394 TTCCATCTCTGACCTTGCTTTCCTGTGCACGCTTCCCATGCTGATAAGGAGTTACGCCAC 92335

Qy      286 TGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAA 345
            || ||  |||||||||||||||||||||||| ||||||||||| ||||||||||| ||||
Db    92334 TGGGAACTGGACCTATGGAGATGTTCTCTGCATAAGCAACCGTTATGTGCTTCATGCCAA 92275

Qy      346 CCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGCTCATGAA 405
            |||||||||||||||||||| |||||||||||||||||||| ||||||||||||||||||||||||||
Db    92274 CCTCTACACCAGCATCCTTTTCCTCACTTTCATTAGCATAGACCGATATCTGCTCATGAA 92215

Qy      406 GTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCTCGCTGGC 465
            || ||||||||||||||||| ||||||||||||||||||||||||||||||||| |||||
Db    92214 GTTCCCTTTCCGAGAACACATTCTACAAAAGAAGGAATTTGCCATTTTAATCTCCCTGGC 92155

Qy      466 TGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATTCTGTCCC 525
            |||||||| ||||||||||||||||||||||| |||||||| || |||| ||| |||
Db    92154 TGTCTGGGTCTTAGTGACCTTAGAAGTTCTACCTATGCTCACGTTTATCACTTCCACCCC 92095

Qy      526 AAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAACACAATCT 585
            || |||| |||||    | ||| |||||||||||||||||||||||||||| || ||| |||
Db    92094 AATAGAAAAGGGCGACAGCTGTGTCGACTATGCAAGTTCTGGAAACCCTAAATACAGTCT 92035

Qy      586 CATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTGTGATGTGCTT 645
            ||||||||||| |||||||||||||| ||||||||||| |||||||||| ||||| ||||||||
Db    92034 CATTTACAGCCTGTGCCTGACTTTGCTGGGCTTCCTCATTCCTCTGTCTGTAATGTGCTT 91975

Qy      646 CTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAACTGCCCT 705
            ||||||||||| ||||||||||||| |||||| |||||||||||||||| |||||||| ||
Db    91974 CTTCTACTACAAAATGGTAGTCTTCCTAAAGAAGAGGAGCCAGCAGCAGGCAACTGTGCT 91915

Qy      706 GCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTATACTCTT 765
            | ||| ||||||| | |||||||||||||||||||||| || ||||||||||| |||||||
Db    91914 ATCGCTGAACAAACCTCTGCGCCTGGTGGTCCTGGCAGTGGTGATCTTCTCTGTACTCTT 91855

Qy      766 CACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTTGGCCACA 825
            ||||| || |||||||||||||||||| ||||| |||||||| |||||||| ||||||||||
Db    91854 TACACCTTACCATATCATGCGCAATGTGAGGATTGCCTCACGCTTGGATAGCTGGCCACA 91795

Qy      826 AGGATGTACACAGAAGGCCATCAAATCTATATACACTGACACGGCCTCTGGCCTTTCT 885
            ||||||| | |||||||||||||||||    ||||||  |||||  | ||||||||||||||
Db    91794 GGGATGTTCCCAGAAGGCCATCAAATGCTTATACATCCTGACCAGACCTCTGGCCTTTCT 91735
```

```
Qy         886 GAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACAGAGAGATGCT 945
               ||||||||| |||| |||||||||||||||||||| ||||||||||| |||||| |||||
Db       91734 GAACAGTGCTGTCAACCCCATCTTCTACTTCCTTGTGGGAGACCATTTCAGAGACATGCT 91675

Qy         946 GATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACATGAGCTGCTGG 1005
               | ||||||||||| |||||||||||||||||||||||| ||||||||||    ||| ||
Db       91674 GTTTAGTAAGTTGAGACAATACTTCAAGTCCCTTACGTCCTTCAGGCTCTGACCT----A 91619

Qy        1006 ATGCAGGTCTTCACTCAGCCAAAATGAGACACTTGATAAACAGTGCTGTGCAGTTGAGTT 1065
               ||| |||||||||||| ||||| ||| |||| |           || |||||||||||||
Db       91618 ATGTAGGTCTTCACTGAGCCAGAATAAGACTC----------AACTCTGCAGTTGAGTT 91570

Qy        1066 TTAACTAAGTAAACCACCATTTCTAGGCTTTAGC-TTTCCACCATCCTCCAACCCCCAGG 1124
               || || ||||| ||||||| |||||||||||||| || ||||||||||||||||| ||
Db       91569 TTGACCAAGTAGACCACCACCTCTAGGCTTTAGCGTTCCCACCATCCTCCAACCCTGAGT 91510

Qy        1125 GCTGGAGTACAAGCTGGGTCCACATGAATCAGAAG-GCAGCTCTCTGTTCTGATTTTAGG 1183
               ||| ||| |||| |||||   ||||||||||||| ||| | ||||| | ||||||||||
Db       91509 GCTAGAGCACAAACTGGGCACACATGAATCAGAAGAGCAACCATCTGTCCCGATTTTAGG 91450

Qy        1184 TTATACCCAGAGTATGGAAAAAATAAGGCATGAGAAAGCATTGACATCTTCACTTAAGAA 1243
               | |||||||||||||| |||||||||| |||  ||||||||||||||||||||||||| ||||||
Db       91449 CTGTACCCAGAGTATGG-AAAAATGAGGCCCCAGAAAGCATTGACATCTTCACATAAGAA 91391

Qy        1244 CTGAACAAAAGAGAACAAATATTGTCAATGTTTGGACACTTAGGATCTGAAATCTTGGAA 1303
               |||||||||||| ||| || ||||||||| ||||||||||| ||||  |     |||||
Db       91390 CTGAACAAAAGAAAACTGATGTTGTCAATATTTGGACACTTAAGATCCAAGGCGTTGGAG 91331

Qy        1304 ATTTTAAGACCTCTT-TTTCTATCAGTGTAAAAGGAATACAAGATAGCTAGTTGCAAATG 1362
               ||||||||||| |||| ||||||||||||||||||||||| ||| ||||||||| | |
Db       91330 ATTTTAAGACATCTTCTTTCTATCAGTGTAAAAGGAATACGAGACAGCTAGTT-CTGACA 91272

Qy        1363 CTGAATGCATTTCATCATTGGTCAGGTCGATAAGCGTGTTTCTGAAATAGTC----TTAT 1418
               ||||||||||| ||||||||||||| |||||||| |||||||||||||||||    ||||
Db       91271 CTGAATGCATTTTGTCATTGGTCAGCTTGATAAGAATGTTTCTGAAATAGTCTCTATTAT 91212

Qy        1419 TTTTATTCTTGTAATATTAA-AATTTATGTGAAAAATGAATATAATTCAATGTACAACAT 1477
               |||||||||| |||||||||   ||||| |||| ||| || || |||||||||||||||
Db       91211 TTTTATTCTTGCAATATTAACCTTTTATATGAATGGTGAGTAGAACTCAATGTACAACAT 91152

Qy        1478 TAGATTTTCTATTTGAAAATTATATTTCTTGAAAA----AATAACTGCTGTGCCTAAATA 1533
               ||| || |||| ||| || ||||||||||    ||||||||| |||||||||
Db       91151 TAGCAATTATATTCAGAAAGTACATTTCTTGAAAAAATGAATAACTGCAATGCCTAAATA 91092

Qy        1534 AATCAATATA 1543
               |||||| | |
Db       91091 AATCAACACA 91082
```

RESULT 5
AC116149
LOCUS   AC116149   60298 bp   DNA   linear   HTG 25-MAR-2002
DEFINITION Mus musculus clone RP24-540E9, LOW-PASS SEQUENCE SAMPLING.
ACCESSION   AC116149
VERSION   AC116149.1 GI:19703273
KEYWORDS   HTG; HTGS_PHASE0.

```
SOURCE        Mus musculus (house mouse)
  ORGANISM    Mus musculus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE     1  (bases 1 to 60298)
  AUTHORS     Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE       Mus musculus, clone RP24-540E9
  JOURNAL     Unpublished
REFERENCE     2  (bases 1 to 60298)
  AUTHORS     Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
              Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
              Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
              Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
              Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
              Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
              Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
              Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
              Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
              Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
              MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
              McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
              Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
              Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
              Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
              Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
              Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
              Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
              Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
              Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
              Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
              Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE       Direct Submission
  JOURNAL     Submitted (25-MAR-2002) Whitehead Institute/MIT Center for Genome
              Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT       All repeats were identified using RepeatMasker:
              Smit, A.F.A. & Green, P. (1996-1997)
              http://ftp.genome.washington.edu/RM/RepeatMasker.html
              -------------- Genome Center
                 Center: Whitehead Institute/ MIT Center for Genome Research
                 Center code: WIBR
                 Web site: http://www-seq.wi.mit.edu
                 Contact: sequence_submissions@genome.wi.mit.edu
              -------------- Project Information
                 Center project name: L24912
                 Center clone name: 540_E_9
              --------------.
              * NOTE: This record contains 77 individual
              * sequencing reads that have not been assembled into
              * contigs. Runs of N are used to separate the reads
              * and the order in which they appear is completely
              * arbitrary. Low-pass sequence sampling is useful for
              * identifying clones that may be gene-rich and allows
              * overlap relationships among clones to be deduced.
              * However, it should not be assumed that this clone
              * will be sequenced to completion. In the event that
              * the record is updated, the accession number will
              * be preserved.
```

```
*          1       656: contig of 656 bp in length
*        657       756: gap of 100 bp
*        757      1426: contig of 670 bp in length
*       1427      1526: gap of 100 bp
*       1527      2210: contig of 684 bp in length
*       2211      2310: gap of 100 bp
*       2311      2997: contig of 687 bp in length
*       2998      3097: gap of 100 bp
*       3098      3786: contig of 689 bp in length
*       3787      3886: gap of 100 bp
*       3887      4577: contig of 691 bp in length
*       4578      4677: gap of 100 bp
*       4678      5357: contig of 680 bp in length
*       5358      5457: gap of 100 bp
*       5458      6150: contig of 693 bp in length
*       6151      6250: gap of 100 bp
*       6251      6817: contig of 567 bp in length
*       6818      6917: gap of 100 bp
*       6918      7615: contig of 698 bp in length
*       7616      7715: gap of 100 bp
*       7716      8412: contig of 697 bp in length
*       8413      8512: gap of 100 bp
*       8513      9198: contig of 686 bp in length
*       9199      9298: gap of 100 bp
*       9299      9988: contig of 690 bp in length
*       9989     10088: gap of 100 bp
*      10089     10768: contig of 680 bp in length
*      10769     10868: gap of 100 bp
*      10869     11524: contig of 656 bp in length
*      11525     11624: gap of 100 bp
*      11625     12242: contig of 618 bp in length
*      12243     12342: gap of 100 bp
*      12343     13040: contig of 698 bp in length
*      13041     13140: gap of 100 bp
*      13141     13829: contig of 689 bp in length
*      13830     13929: gap of 100 bp
*      13930     14647: contig of 718 bp in length
*      14648     14747: gap of 100 bp
*      14748     15451: contig of 704 bp in length
*      15452     15551: gap of 100 bp
*      15552     16247: contig of 696 bp in length
*      16248     16347: gap of 100 bp
*      16348     17028: contig of 681 bp in length
*      17029     17128: gap of 100 bp
*      17129     17802: contig of 674 bp in length
*      17803     17902: gap of 100 bp
*      17903     18593: contig of 691 bp in length
*      18594     18693: gap of 100 bp
*      18694     19375: contig of 682 bp in length
*      19376     19475: gap of 100 bp
*      19476     20082: contig of 607 bp in length
*      20083     20182: gap of 100 bp
*      20183     20875: contig of 693 bp in length
*      20876     20975: gap of 100 bp
*      20976     21650: contig of 675 bp in length
*      21651     21750: gap of 100 bp
*      21751     22427: contig of 677 bp in length
```

```
*      22428      22527: gap of 100 bp
*      22528      23238: contig of 711 bp in length
*      23239      23338: gap of 100 bp
*      23339      24028: contig of 690 bp in length
*      24029      24128: gap of 100 bp
*      24129      24803: contig of 675 bp in length
*      24804      24903: gap of 100 bp
*      24904      25603: contig of 700 bp in length
*      25604      25703: gap of 100 bp
*      25704      26357: contig of 654 bp in length
*      26358      26457: gap of 100 bp
*      26458      27140: contig of 683 bp in length
*      27141      27240: gap of 100 bp
*      27241      27946: contig of 706 bp in length
*      27947      28046: gap of 100 bp
*      28047      28734: contig of 688 bp in length
*      28735      28834: gap of 100 bp
*      28835      29536: contig of 702 bp in length
*      29537      29636: gap of 100 bp
*      29637      30324: contig of 688 bp in length
*      30325      30424: gap of 100 bp
*      30425      31130: contig of 706 bp in length
*      31131      31230: gap of 100 bp
*      31231      31910: contig of 680 bp in length
*      31911      32010: gap of 100 bp
*      32011      32691: contig of 681 bp in length
*      32692      32791: gap of 100 bp
*      32792      33482: contig of 691 bp in length
*      33483      33582: gap of 100 bp
*      33583      34274: contig of 692 bp in length
*      34275      34374: gap of 100 bp
*      34375      35081: contig of 707 bp in length
*      35082      35181: gap of 100 bp
*      35182      35861: contig of 680 bp in length
*      35862      35961: gap of 100 bp
*      35962      36660: contig of 699 bp in length
*      36661      36760: gap of 100 bp
*      36761      37447: contig of 687 bp in length
*      37448      37547: gap of 100 bp
*      37548      38243: contig of 696 bp in length
*      38244      38343: gap of 100 bp
*      38344      39034: contig of 691 bp in length
*      39035      39134: gap of 100 bp
*      39135      39813: contig of 679 bp in length
*      39814      39913: gap of 100 bp
*      39914      40597: contig of 684 bp in length
*      40598      40697: gap of 100 bp
*      40698      41392: contig of 695 bp in length
*      41393      41492: gap of 100 bp
*      41493      42190: contig of 698 bp in length
*      42191      42290: gap of 100 bp
*      42291      42967: contig of 677 bp in length
*      42968      43067: gap of 100 bp
*      43068      43736: contig of 669 bp in length
*      43737      43836: gap of 100 bp
*      43837      44525: contig of 689 bp in length
*      44526      44625: gap of 100 bp
```

```
        *      44626     45306: contig of 681 bp in length
        *      45307     45406: gap of 100 bp
        *      45407     46111: contig of 705 bp in length
        *      46112     46211: gap of 100 bp
        *      46212     46848: contig of 637 bp in length
        *      46849     46948: gap of 100 bp
        *      46949     47639: contig of 691 bp in length
        *      47640     47739: gap of 100 bp
        *      47740     48431: contig of 692 bp in length
        *      48432     48531: gap of 100 bp
        *      48532     49221: contig of 690 bp in length
        *      49222     49321: gap of 100 bp
        *      49322     50017: contig of 696 bp in length
        *      50018     50117: gap of 100 bp
        *      50118     50799: contig of 682 bp in length
        *      50800     50899: gap of 100 bp
        *      50900     51583: contig of 684 bp in length
        *      51584     51683: gap of 100 bp
        *      51684     52384: contig of 701 bp in length
        *      52385     52484: gap of 100 bp
        *      52485     53167: contig of 683 bp in length
        *      53168     53267: gap of 100 bp
        *      53268     53966: contig of 699 bp in length
        *      53967     54066: gap of 100 bp
```

```
    Query Match              41.9%;   Score 645.8;   DB 2;    Length 60298;
    Best Local Similarity    84.0%;   Pred. No. 2.3e-133;
    Matches   673;  Conservative    0;  Mismatches  127;   Indels    1;  Gaps    1;
```

```
Qy        51 AGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCT 110
             |||   | || |     |       |||| |       |  |||||||||||||||||||||||||||||||||
Db      3890 AGATCTGATATCTCGCCCTGTGGTGGAATTCTCAGGCTATCTTGAATAAGTACTACCTCT 3949

Qy       111 CTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGTGTTCG 170
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      3950 CTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGTGTTTG 4009

Qy       171 GCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCTTTCCA 230
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4010 GCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCTTTCCA 4069

Qy       231 TCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAATGATA 290
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4070 TCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAATGATA 4129

Qy       291 AGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAACCTCT 350
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4130 AGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAACCTCT 4189

Qy       351 ACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGCTCATGAAGTACC 410
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4190 ACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGCTCATGAAGTACC 4249

Qy       411 CTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCTCGCTGGCTGTCT 470
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4250 CTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCTCGCTGGCTGTCT 4309
```

```
Qy      471 GGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATTCTGTCCCAAAG 530
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4310 GGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATTCTGTCCCAAAG 4369

Qy      531 AAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAACACAATCTCATTT 590
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4370 AAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAACACAATCTCATTT 4429

Qy      591 ACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGATGTGCTTCTTCT 650
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4430 ACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGATGTGCTTCTTCT 4489

Qy      651 ACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAACTGCCCTGCCAC 710
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4490 ACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAACTGCCCTGCCAC 4549

Qy      711 TGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTATACTCTTCACAC 770
            |||||  |||||  |||||||||| |||||||
Db     4550 TGGAC-AACCCAAACGCCTGGGGGTCCTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4608

Qy      771 CCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTTGGCCACAAGGAT 830

Db     4609 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4668

Qy      831 GTACACAGAAGGCCATCAAAT 851
                ||   |||   ||
Db     4669 NNNNNNNNNNCGGAGATCTGAT 4689


RESULT 6
AC116149/c
LOCUS       AC116149              60298 bp    DNA     linear   HTG 25-MAR-2002
DEFINITION  Mus musculus clone RP24-540E9, LOW-PASS SEQUENCE SAMPLING.
ACCESSION   AC116149
VERSION     AC116149.1  GI:19703273
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 60298)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus, clone RP24-540E9
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 60298)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
            Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
```

Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (25-MAR-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
  COMMENT   All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L24912
                Center clone name: 540_E_9
            --------------.

            * NOTE: This record contains 77 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *   1   656: contig of 656 bp in length
            *   657   756: gap of 100 bp
            *   757   1426: contig of 670 bp in length
            *   1427   1526: gap of 100 bp
            *   1527   2210: contig of 684 bp in length
            *   2211   2310: gap of 100 bp
            *   2311   2997: contig of 687 bp in length
            *   2998   3097: gap of 100 bp
            *   3098   3786: contig of 689 bp in length
            *   3787   3886: gap of 100 bp
            *   3887   4577: contig of 691 bp in length
            *   4578   4677: gap of 100 bp
            *   4678   5357: contig of 680 bp in length
            *   5358   5457: gap of 100 bp
            *   5458   6150: contig of 693 bp in length
            *   6151   6250: gap of 100 bp
            *   6251   6817: contig of 567 bp in length
            *   6818   6917: gap of 100 bp
            *   6918   7615: contig of 698 bp in length
            *   7616   7715: gap of 100 bp
            *   7716   8412: contig of 697 bp in length

```
*      8413      8512: gap of 100 bp
*      8513      9198: contig of 686 bp in length
*      9199      9298: gap of 100 bp
*      9299      9988: contig of 690 bp in length
*      9989     10088: gap of 100 bp
*     10089     10768: contig of 680 bp in length
*     10769     10868: gap of 100 bp
*     10869     11524: contig of 656 bp in length
*     11525     11624: gap of 100 bp
*     11625     12242: contig of 618 bp in length
*     12243     12342: gap of 100 bp
*     12343     13040: contig of 698 bp in length
*     13041     13140: gap of 100 bp
*     13141     13829: contig of 689 bp in length
*     13830     13929: gap of 100 bp
*     13930     14647: contig of 718 bp in length
*     14648     14747: gap of 100 bp
*     14748     15451: contig of 704 bp in length
*     15452     15551: gap of 100 bp
*     15552     16247: contig of 696 bp in length
*     16248     16347: gap of 100 bp
*     16348     17028: contig of 681 bp in length
*     17029     17128: gap of 100 bp
*     17129     17802: contig of 674 bp in length
*     17803     17902: gap of 100 bp
*     17903     18593: contig of 691 bp in length
*     18594     18693: gap of 100 bp
*     18694     19375: contig of 682 bp in length
*     19376     19475: gap of 100 bp
*     19476     20082: contig of 607 bp in length
*     20083     20182: gap of 100 bp
*     20183     20875: contig of 693 bp in length
*     20876     20975: gap of 100 bp
*     20976     21650: contig of 675 bp in length
*     21651     21750: gap of 100 bp
*     21751     22427: contig of 677 bp in length
*     22428     22527: gap of 100 bp
*     22528     23238: contig of 711 bp in length
*     23239     23338: gap of 100 bp
*     23339     24028: contig of 690 bp in length
*     24029     24128: gap of 100 bp
*     24129     24803: contig of 675 bp in length
*     24804     24903: gap of 100 bp
*     24904     25603: contig of 700 bp in length
*     25604     25703: gap of 100 bp
*     25704     26357: contig of 654 bp in length
*     26358     26457: gap of 100 bp
*     26458     27140: contig of 683 bp in length
*     27141     27240: gap of 100 bp
*     27241     27946: contig of 706 bp in length
*     27947     28046: gap of 100 bp
*     28047     28734: contig of 688 bp in length
*     28735     28834: gap of 100 bp
*     28835     29536: contig of 702 bp in length
*     29537     29636: gap of 100 bp
*     29637     30324: contig of 688 bp in length
*     30325     30424: gap of 100 bp
```

```
*      30425      31130: contig of 706 bp in length
*      31131      31230: gap of 100 bp
*      31231      31910: contig of 680 bp in length
*      31911      32010: gap of 100 bp
*      32011      32691: contig of 681 bp in length
*      32692      32791: gap of 100 bp
*      32792      33482: contig of 691 bp in length
*      33483      33582: gap of 100 bp
*      33583      34274: contig of 692 bp in length
*      34275      34374: gap of 100 bp
*      34375      35081: contig of 707 bp in length
*      35082      35181: gap of 100 bp
*      35182      35861: contig of 680 bp in length
*      35862      35961: gap of 100 bp
*      35962      36660: contig of 699 bp in length
*      36661      36760: gap of 100 bp
*      36761      37447: contig of 687 bp in length
*      37448      37547: gap of 100 bp
*      37548      38243: contig of 696 bp in length
*      38244      38343: gap of 100 bp
*      38344      39034: contig of 691 bp in length
*      39035      39134: gap of 100 bp
*      39135      39813: contig of 679 bp in length
*      39814      39913: gap of 100 bp
*      39914      40597: contig of 684 bp in length
*      40598      40697: gap of 100 bp
*      40698      41392: contig of 695 bp in length
*      41393      41492: gap of 100 bp
*      41493      42190: contig of 698 bp in length
*      42191      42290: gap of 100 bp
*      42291      42967: contig of 677 bp in length
*      42968      43067: gap of 100 bp
*      43068      43736: contig of 669 bp in length
*      43737      43836: gap of 100 bp
*      43837      44525: contig of 689 bp in length
*      44526      44625: gap of 100 bp
*      44626      45306: contig of 681 bp in length
*      45307      45406: gap of 100 bp
*      45407      46111: contig of 705 bp in length
*      46112      46211: gap of 100 bp
*      46212      46848: contig of 637 bp in length
*      46849      46948: gap of 100 bp
*      46949      47639: contig of 691 bp in length
*      47640      47739: gap of 100 bp
*      47740      48431: contig of 692 bp in length
*      48432      48531: gap of 100 bp
*      48532      49221: contig of 690 bp in length
*      49222      49321: gap of 100 bp
*      49322      50017: contig of 696 bp in length
*      50018      50117: gap of 100 bp
*      50118      50799: contig of 682 bp in length
*      50800      50899: gap of 100 bp
*      50900      51583: contig of 684 bp in length
*      51584      51683: gap of 100 bp
*      51684      52384: contig of 701 bp in length
*      52385      52484: gap of 100 bp
*      52485      53167: contig of 683 bp in length
```

```
        *      53168    53267: gap of 100 bp
        *      53268    53966: contig of 699 bp in length
        *      53967    54066: gap of 100 bp

   Query Match            41.1%;  Score 633.6;  DB 2;  Length 60298;
   Best Local Similarity  97.6%;  Pred. No. 1.2e-130;
    Matches  664;  Conservative   0;  Mismatches  14;  Indels   2;  Gaps   2;

Qy      379 TAGCATGGACCGATATCTGCTCATGAAGTACCCTTTCCGAG-AACACTTTCTACAAAA-G 436
            |||||||||||||||||||||||||||||||||||||||| |||||| ||||||||||||| |
Db    36659 TAGCATGGACCGATATCTGCTCATGAAGTACCCTTCCCGAGAAACACTTTCTACAAAANG 36600

Qy      437 AAGGAATTTGCCATTTTAATCTCGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTA 496
            ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db    36599 AAGGAATTTGCCATTTTAATCTCGCTGGCTGTCTGGGCTTTAGTGACCTTAGAAGTTCTA 36540

Qy      497 CCCATGCTCACTTTCATCAATTCTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTAT 556
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    36539 CCCATGCTCACTTTCATCAATTCTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTAT 36480

Qy      557 GCAAGTTCTGGAAACCCTGAACACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGC 616
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    36479 GCAAGTTCTGGAAACCCTGAACACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGC 36420

Qy      617 TTCCTAATTCCTCTCTCTGTGATGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAG 676
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    36419 TTCCTAATTCCTCTCTCTGTGATGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAG 36360

Qy      677 AGGAGGAGCCAGCAGCAAGCAACTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTC 736
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    36359 AGGAGGAGCCAGCAGCAAGCAACTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTC 36300

Qy      737 CTGGCGGTTGTGATCTTCTCTATACTCTTCACACCCTATCATATCATGCGCAATTTGAGG 796
            ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    36299 CTGGCAGTTGTGATCTTCTCTATACTCTTCACACCCTATCATATCATGCGCAATTTGAGG 36240

Qy      797 ATCGCCTCACGCCTGGATAGTTGGCCACAAGGATGTACACAGAAGGCCATCAAATCTATA 856
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    36239 ATCGCCTCACGCCTGGATAGTTGGCCACAAGGATGTACACAGAAGGCCATCAAATCTATA 36180

Qy      857 TACACACTGACACGGCCTCTGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTC 916
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    36179 TACACACTGACACGGCCTCTGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTC 36120

Qy      917 CTCATGGGAGACCATTACAGAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCC 976
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    36119 CTCATGGGAGACCATTACAGAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCC 36060

Qy      977 CTTACATCCTTCAGGACATGAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAATGAGACA 1036
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    36059 CTTACATCCTTCAGGACATGAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAATGAGACA 36000

Qy     1037 CTTGATAAACAGTGCTGTGC 1056
            ||    |  ||   | | ||
Db    35999 CTGAGAATCCACCACAGGGC 35980
```

RESULT 7
AC110839/c
LOCUS   AC110839   326606 bp   DNA   linear   HTG 11-OCT-2002
DEFINITION Rattus norvegicus clone CH230-208A12, *** SEQUENCING IN PROGRESS
            ***, 25 unordered pieces.
ACCESSION   AC110839
VERSION   AC110839.4 GI:23820318
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE   Rattus norvegicus (Norway rat)
  ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 326606)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,

          Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
          Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
          Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
          Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
          Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
          Weinstock,G. and Gibbs,R.A.
  TITLE   Direct Submission
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 326606)
  AUTHORS   Worley,K.C.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-FEB-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3 (bases 1 to 326606)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE   Direct Submission
  JOURNAL   Submitted (11-OCT-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT   On Oct 11, 2002 this sequence version replaced gi:21739250.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GRKD
                Center clone name: CH230-208A12
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 242752 bases at least Q40
                Consensus quality: 250821 bases at least Q30
                Consensus quality: 254983 bases at least Q20
                Estimated insert size: 244968; sum-of-contigs estimation
                Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
            ----
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 25 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence

```
* as soon as it is available and the accession number will
* be preserved.
*         1     10356: contig of 10356 bp in length
*     10357     10456: gap of unknown length
*     10457     15819: contig of 5363 bp in length
*     15820     15919: gap of unknown length
*     15920    245368: contig of 229449 bp in length
*    245369    245468: gap of unknown length
*    245469    272041: contig of 26573 bp in length
*    272042    272141: gap of unknown length
*    272142    276368: contig of 4227 bp in length
*    276369    276468: gap of unknown length
*    276469    282159: contig of 5691 bp in length
*    282160    282259: gap of unknown length
*    282260    283432: contig of 1173 bp in length
*    283433    283532: gap of unknown length
*    283533    284633: contig of 1101 bp in length
*    284634    284733: gap of unknown length
*    284734    285764: contig of 1031 bp in length
*    285765    285864: gap of unknown length
*    285865    287082: contig of 1218 bp in length
*    287083    287182: gap of unknown length
*    287183    288399: contig of 1217 bp in length
*    288400    288499: gap of unknown length
*    288500    289828: contig of 1329 bp in length
*    289829    289928: gap of unknown length
*    289929    291274: contig of 1346 bp in length
*    291275    291374: gap of unknown length
*    291375    293018: contig of 1644 bp in length
*    293019    293118: gap of unknown length
*    293119    294732: contig of 1614 bp in length
*    294733    294832: gap of unknown length
*    294833    296078: contig of 1246 bp in length
*    296079    296178: gap of unknown length
*    296179    297942: contig of 1764 bp in length
*    297943    298042: gap of unknown length
*    298043    299812: contig of 1770 bp in length
*    299813    299912: gap of unknown length
*    299913    301595: contig of 1683 bp in length
*    301596    301695: gap of unknown length
*    301696    304787: contig of 3092 bp in length
*    304788    304887: gap of unknown length
*    304888    306249: contig of 1362 bp in length
*    306250    306349: gap of unknown length
*    306350    307801: contig of 1452 bp in length
*    307802    307901: gap of unknown length
*    307902    309454: contig of 1553 bp in length
*    309455    309554: gap of unknown length
*    309555    314110: contig of 4556 bp in length
*    314111    314210: gap of unknown length
*    314211    326606: contig of 12396 bp in length.
FEATURES             Location/Qualifiers
     source          1. .326606
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-208A12"
```

```
       misc_feature      1. .1742
                         /note="wgs_end_extension
                         clone_end:Sp6"
       misc_feature      complement(4245. .5082)
                         /note="clone_boundary
                         clone_end:Sp6
                         site:EcoRI
                         end_sequence:RWBKN06TVB"
       misc_feature      10457. .12850
                         /note="wgs_contig"
       misc_feature      15920. .16991
                         /note="wgs_contig"
       misc_feature      complement(220129. .221101)
                         /note="clone_boundary
                         clone_end:T7
                         site:EcoRI
                         end_sequence:RWBKN06TJB"
       misc_feature      241580. .242749
                         /note="wgs_end_extension
                         clone_end:T7"
       misc_feature      243833. .245368
                         /note="wgs_end_extension
                         clone_end:T7"
ORIGIN

   Query Match              39.9%;   Score 615.8;   DB 2;   Length 326606;
   Best Local Similarity    89.0%;   Pred. No. 1.2e-126;
   Matches  665;  Conservative    0;  Mismatches   82;   Indels    0;   Gaps    0;

Qy       46 GGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTA 105
            |||||||||||||||||||||||| |||||| ||||| |||| ||| ||||| ||||||||
Db   242326 GGCACAGAATTTATCTTGTGAAAATTGGCTGGCATTAGAGAATATTTTGAAAAAGTACTA
     242267

Qy      106 CCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGT 165
            ||||||||||||||||     |||||||||||| ||||| ||||||| ||| |||| |||||
Db   242266 CCTCTCTGCATTTTATGGGATCGAGTTCATTGTTGGAATGCTTGGCAATTTCACCGTGGT
     242207

Qy      166 GTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCT 225
            ||||||||||||||||||||||||||||||||||||||||| ||||| ||||||||| ||||||
Db   242206 GTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGTAGCAACGTCTATCTCTTCAACCT
     242147

Qy      226 TTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAA 285
            |||||||||||| |||||||||||||||||| |||||||| ||||||| |||||| ||||
Db   242146 TTCCATCTCTGACCTTGCTTTCCTGTGCACGCTTCCCATGCTGATAAGGAGTTACGCCAC
     242087

Qy      286 TGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAA 345
            ||  ||  ||||||||||||||||||||||| ||||||||||| |||||||||| ||||
Db   242086 TGGGAACTGGACCTATGGAGATGTTCTCTGCATAAGCAACCGTTATGTGCTTCATGCCAA
     242027

Qy      346 CCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGCTCATGAA 405
            |||||||||||||||||||| ||||||||||||||||||||| |||||||||||||||||||
```

```
Db       242026 CCTCTACACCAGCATCCTTTTCCTCACTTTCATTAGCATAGACCGATATCTGCTCATGAA
         241967

Qy          406 GTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCTCGCTGGC 465
                || ||||||||||||||||| |||||||||||||||||||||||||||||||||| |||||
Db       241966 GTTCCCTTTCCGAGAACACATTCTACAAAGAAGGAATTTGCCATTTTAATCTCCCTGGC
         241907

Qy          466 TGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATTCTGTCCC 525
                |||||||| |||||||||||||||||||||||| ||||||||| || |||| |||   |||
Db       241906 TGTCTGGGTCTTAGTGACCTTAGAAGTTCTACCTATGCTCACGTTTATCACTTCCACCCC
         241847

Qy          526 AAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAACACAATCT 585
                || |||| |||||     | ||| ||||||||||||||||||||||||| || ||| |||
Db       241846 AATAGAAAGGGCGACAGCTGTGTCGACTATGCAAGTTCTGGAAACCCTAAATACAGTCT
         241787

Qy          586 CATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGATGTGCTT 645
                |||||||||||| |||||||||||||| ||||||||| ||||||||| ||||| ||||||||
Db       241786 CATTTACAGCCTGTGCCTGACTTTGCTGGGCTTCCTCATTCCTCTGTCTGTAATGTGCTT
         241727

Qy          646 CTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAACTGCCCT 705
                |||||||||||| |||||||||||||| |||||||| |||||||||||||| |||||| ||
Db       241726 CTTCTACTACAAAATGGTAGTCTTCCTAAAGAAGAGGAGCCAGCAGCAGGCAACTGTGCT
         241667

Qy          706 GCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTATACTCTT 765
                | ||| |||||||| | |||||||||||||||||| || |||||||||||| |||||||
Db       241666 ATCGCTGAACAAACCTCTGCGCCTGGTGGTCCTGGCAGTGGTGATCTTCTCTGTACTCTT
         241607

Qy          766 CACACCCTATCATATCATGCGCAATTT 792
                ||||| || |||||||||||||||| |
Db       241606 TACACCTTACCATATCATGCGCAATGT 241580


RESULT 8
AF247785
LOCUS       AF247785                1325 bp    mRNA    linear   PRI 26-MAR-2002
DEFINITION  Homo sapiens P2Y purinoceptor 1 mRNA, complete cds.
ACCESSION   AF247785
VERSION     AF247785.1  GI:19716154
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1325)
  AUTHORS   Zhang,W., Li,N., Wan,T. and Cao,X.
  TITLE     Human P2Y purinoceptor 1
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1325)
  AUTHORS   Zhang,W., Li,N., Wan,T. and Cao,X.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAR-2000) Department of Immunology, Second Military
```

```
FEATURES                Location/Qualifiers
     source             1. .1325
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
     CDS                69. .1073
                        /codon_start=1
                        /product="P2Y purinoceptor 1"
                        /protein_id="AAL95690.1"
                        /db_xref="GI:19716155"
                        /translation="MLGIMAWNATCKNWLAAEAALEKYYLSIFYGIEFVVGVLGNTIV
                        VYGYIFSLKNWNSSNIYLFNLSVSDLAFLCTLPMLIRSYANGNWIYGDVLCISNRYVL
                        HANLYTSILFLTFISIDRYLIIKYPFREHLLQKKEFAILISLAIWVLVTLELLPILPL
                        INPVITDNGTTCNDFASSGDPNYNLIYSMCLTLLGFLIPLFVMCFFYYKIALFLKQRN
                        RQVATALPLEKPLNLVIMAVVIFSVLFTPYHVMRNVRIASRLGSWKQYQCTQVVINSF
                        YIVTRPLAFLNSVINPVFYFLLGDHFRDMLMNQLRHNFKSLTSFSRWAHELLLSFREK
                        "
ORIGIN

   Query Match          38.4%;   Score 592.4;   DB 9;   Length 1325;
   Best Local Similarity  75.3%;   Pred. No. 1.6e-121;
   Matches  764;  Conservative    0;  Mismatches  246;  Indels    4;  Gaps    2;

Qy       39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
            | |   |||||||  ||||  |  ||||  |  || |||  |||||  ||||||||  |  ||  | |
Db       76 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 135

Qy       99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
            |||||||||||| ||     |||||||  || ||||||  || | ||| | ||||| |||   ||
Db      136 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 195

Qy      159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
            ||| || | ||||||| |||||| |||||||||||||||||||||||||| ||| | |||||| |
Db      196 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 255

Qy      219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
            |||||||  ||  ||||||||||| |||||  ||||||||||||| ||||| |||||||| |||||
Db      256 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 315

Qy      279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
            |||||||||  ||  ||| |||||||| || |||||  ||||||||||||||||||||||||||||
Db      316 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 375

Qy      339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
            |  ||||||||| ||||||||| |||||| ||||| || ||||| || ||||| ||
Db      376 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 435

Qy      399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
            | || |||||| ||||||||||||||||| |||| ||||||||| || ||||| |||||||||||
Db      436 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAGAAAGAGTTTGCTATTTTAATCT 495

Qy      459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
            | |||| | |||| |||||| ||||||||| | ||||||||| ||  |  | || |||
Db      496 CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 555
```

```
Qy    519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
          ||||   ||  || | |||| | ||| | || | |||||||||||||| |||| |
Db    556 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 615

Qy    579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
          |||| |||||||||||| | || || || ||||||| ||||| ||||||||| | |||||
Db    616 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTGTGA 675

Qy    639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
          |||| ||||| || ||||||||| |     ||||| ||||| |||| | |||
Db    676 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 735

Qy    699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
          |||| |||||| || || || || | | |||| || |||| || || |||||||||
Db    736 CTGCTCTGCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 795

Qy    759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
          | || ||  |||||||||| || |||||| ||| ||||||||||  |||||||||| ||||
Db    796 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 855

Qy    819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
          |   ||  |    || || ||| | |||||| || | |||| |||||||||||
Db    856 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT 915

Qy    876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
          ||||||||||||||||| ||||||| || ||||||| || ||  ||||||| || | ||
Db    916 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 975

Qy    936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
          | || ||||||||| | | | | ||||| ||||||| |||||||||||||| || | ||
Db    976 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 1035

Qy    996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
          | ||| || | |    |||| |||| |||| |||    |||| ||||||
Db   1036 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAAGTGAGGGGCTTGTGAAACAG 1089
```

RESULT 9
AX549281
```
LOCUS       AX549281               1380 bp    DNA     linear   PAT 26-NOV-2002
DEFINITION  Sequence 566 from Patent WO02061087.
ACCESSION   AX549281
VERSION     AX549281.1  GI:25813951
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Burmer,G.C., Roush,C.L. and Brown,J.P.
  TITLE     Antigenic peptides, such as for G protein-coupled receptors
            (GPCRs), antibodies thereto, and systems for identifying such
            antigenic peptides
  JOURNAL   Patent: WO 02061087-A 566 08-AUG-2002;
            Lifespan Biosciences, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..1380
```

ORIGIN

```
Query Match              38.4%;  Score 592.4;  DB 6;   Length 1380;
Best Local Similarity    75.3%;  Pred. No. 1.6e-121;
Matches  764;  Conservative    0;  Mismatches 246;  Indels    4;  Gaps    2;
```

```
Qy        39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
             | |  ||||||  ||||  | ||||  | || ||| ||||| ||||||||  | || | |
Db        50 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 109

Qy        99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
             |||||||||| ||     ||||||| || ||||||| || | ||| | ||||| |||  ||
Db       110 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 169

Qy       159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
             ||| || | ||||||| ||||||     |||||||||||||||||| ||| | |||||| |
Db       170 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 229

Qy       219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
             ||||||| ||  |||||||||| |||||| ||||| ||||||||| ||||| ||||||| |||||
Db       230 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 289

Qy       279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
             ||||||||| || ||| |||||||| || ||||| |||||||||||||||||||||||||||||||
Db       290 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 349

Qy       339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
             |  ||||||||| |||||||||| |||||| ||||||||| || ||||| || ||||| ||
Db       350 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 409

Qy       399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
             | || ||||| ||||||||||||||||| |||| ||||||||| || ||||| ||||||||||
Db       410 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCT 469

Qy       459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
             | ||||  |  |||| ||||| |||||||| | |||||||| ||  |  | || |||
Db       470 CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 529

Qy       519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
             ||||    ||  ||  | ||||  | |||  ||  | |||||||||||||||| |||| |
Db       530 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 589

Qy       579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
             ||||  |||||||||| | || |  || || ||||||| |||| ||||| |||||||| |||||
Db       590 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 649

Qy       639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
             |||| |||||| || |||||||||| | |||||||  | ||||| ||||| |||    |||| || |
Db       650 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 709

Qy       699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
             |||| |||||| || || || || || |    | |||| ||  ||||| || || ||||||||||
Db       710 CTGCTCTGCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 769
```

```
Qy    759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
          | || ||  ||||||||||| |||||||| ||| |||||||||| ||||||||| ||||
Db    770 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 829

Qy    819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
          |   || |    || || ||| | |||||| || | |||    |||||||||||
Db    830 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT 889

Qy    876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
          |||||||||||||||||| |||||| || ||||||| || || ||||||| || | ||
Db    890 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 949

Qy    936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
          | || ||||||||| | | | | ||||| ||||||| ||||||||||||||| || | ||
Db    950 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 1009

Qy    996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
          | ||| || | |    |||| |||| |||| |||| |||| ||||||
Db   1010 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAAGTGAGGGGCTTGTGAAACAG 1063
```

RESULT 10
AX780453

```
LOCUS       AX780453                1380 bp    DNA        linear    PAT 14-JUL-2003
DEFINITION  Sequence 2610 from Patent WO03039443.
ACCESSION   AX780453
VERSION     AX780453.1  GI:32697447
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Haferlach,T., Schoch,C., Kern,W., Kohlmann,A., Schnittger,S.,
            Dugas,M., Eils,R., Brors,B. and Mergenthaler,S.
  TITLE     Novel genetic markers for leukemias
  JOURNAL   Patent: WO 03039443-A 2610 15-MAY-2003;
            Deutsches Krebsforschungszentrum (DE) ;
            Ludwig-Maximilian-Universitaet Muenchen (DE) ; Haferlach, Torsten,
            PD Dr. Dr. (DE) ; Schoch, Claudia (DE) ; Kern, Wolfgang (DE)
FEATURES            Location/Qualifiers
     source         1. .1380
                    /organism="Homo sapiens"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:9606"
ORIGIN
```

```
Query Match            38.4%;   Score 592.4;  DB 6;  Length 1380;
Best Local Similarity  75.3%;   Pred. No. 1.6e-121;
Matches  764; Conservative    0;  Mismatches  246;  Indels    4;  Gaps    2;
```

```
Qy     39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
          | |  |||||| |||| |||| | |||| | || ||| ||||| ||||||||| | || |
Db     50 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 109

Qy     99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
          |||||||||| ||        |||||||| || ||||||| || | ||| | ||||| ||
```

```
Db   110 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 169

Qy   159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
         ||| || | ||||||| ||||||         ||||||||||||||||||| ||| | ||||| |
Db   170 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 229

Qy   219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
         |||||| || |||||||||| ||||| |||||||||| |||||| ||||||| |||||
Db   230 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 289

Qy   279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
         |||||||| || ||| ||||||||| || ||||| ||||||||||||||||||||||||||
Db   290 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 349

Qy   339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
         | ||||||||| ||||||||| ||||| ||||||||| || ||||| || ||||| ||
Db   350 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 409

Qy   399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
         | || |||||| ||||||||||||||||||| |||| |||||||| || ||||| |||||||||||
Db   410 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCT 469

Qy   459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
         | |||| | |||| ||||| ||||||||| | |||||||| || | | || ||| |||
Db   470 CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 529

Qy   519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
         ||||    || || | |||| | ||| | || | |||||||||||||||||| ||||| |
Db   530 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 589

Qy   579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
         |||| |||||||||||| | || || || |||||||| ||||| ||||||||| |||||
Db   590 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 649

Qy   639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
         |||| ||||| || |||||||| | ||||| ||||| ||||| ||||| |||| || |
Db   650 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 709

Qy   699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
         |||| |||||| || || || || | | | |||| || ||||| || || ||||||||||
Db   710 CTGCTCTGCCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 769

Qy   759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
         | || || ||||||||| ||||||| ||||| |||||||||||| |||||||||| ||||
Db   770 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 829

Qy   819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
         | || | || || || | ||||||| || | |||| ||||||||||
Db   830 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT 889

Qy   876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
         ||||||||||||||||||||||| ||||||| || |||||||| || || ||||||| || | ||
Db   890 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 949

Qy   936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
         | || |||||||| | | | | ||||| ||||||| |||||||||||||||||| || | ||
Db   950 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 1009
```

```
Qy      996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
            |  |||   || |    |       |||| ||||  |||| ||||   |||| ||||||
Db     1010 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAAGTGAGGGGCTTGTGAAACAG 1063
```

RESULT 11
AF348078

```
LOCUS           AF348078                1380 bp    mRNA    linear   PRI 03-APR-2001
DEFINITION      Homo sapiens G-protein coupled receptor 91 (GPR91) mRNA, complete
                cds.
ACCESSION       AF348078
VERSION         AF348078.1  GI:13517982
KEYWORDS        .
SOURCE          Homo sapiens (human)
  ORGANISM      Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE       1  (bases 1 to 1380)
  AUTHORS       Wittenberger,T., Schaller,H.C. and Hellebrand,S.
  TITLE         An expressed sequence tag (EST) data mining strategy succeeding in
                the discovery of new G-protein coupled receptors
  JOURNAL       J. Mol. Biol. 307 (3), 799-813 (2001)
  MEDLINE       21172992
   PUBMED       11273702
REFERENCE       2  (bases 1 to 1380)
  AUTHORS       Wittenberger,T., Schaller,C.H. and Hellebrand,S.
  TITLE         Direct Submission
  JOURNAL       Submitted (08-FEB-2001) ZMNH, Institut fur
                Entwicklungsneurobiologie, Martinistr. 52, Hamburg 20246, Germany
FEATURES             Location/Qualifiers
     source          1..1380
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /map="3q24-q25.1"
     gene            1..1380
                     /gene="GPR91"
     CDS             55..1047
                     /gene="GPR91"
                     /note="orphan receptor"
                     /codon_start=1
                     /product="G-protein coupled receptor 91"
                     /protein_id="AAK29080.1"
                     /db_xref="GI:13517983"
                     /translation="MAWNATCKNWLAAEAALEKYYLSIFYGIEFVVGVLGNTIVVYGY
                     IFSLKNWNSSNIYLFNLSVSDLAFLCTLPMLIRSYANGNWIYGDVLCISNRYVLHANL
                     YTSILFLTFISIDRYLIIKYPFREHLLQKKEFAILISLAIWVLVTLELLPILPLINPV
                     ITDNGTTCNDFASSGDPNYNLIYSMCLTLLGFLIPLFVMCFFYYKIALFLKQRNRQVA
                     TALPLEKPLNLVIMAVVIFSVLFTPYHVMRNVRIASRLGSWKQYQCTQVVINSFYIVT
                     RPLAFLNSVINPVFYFLLGDHFRDMLMNQLRHNFKSLTSFSRWAHELLLSFREK"
ORIGIN
```

```
Query Match            38.4%;  Score 592.4;  DB 9;  Length 1380;
Best Local Similarity  75.3%;  Pred. No. 1.6e-121;
  Matches 764;  Conservative    0;  Mismatches 246;  Indels    4;  Gaps    2;
```

```
Qy    39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA  98
          | |   |||||||    ||||   | ||||   | ||  ||| ||||| |||||||||  | | ||  |
Db    50 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 109

Qy    99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
          |||||||||||  ||       |||||||   ||  ||||||| ||  | ||| | ||||| |||   ||
Db   110 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 169

Qy   159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
          |||  || | |||||||| |||||| |||||      |||||||||||||||||||| ||| | ||||| |
Db   170 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 229

Qy   219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
          ||||||| ||  |||||||||| |||||| ||||| ||||||||||||| |||||| |||||| |||||
Db   230 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 289

Qy   279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
          |||||||||   ||   |||   ||||||||| | |||||  ||||| |||||||||||||||||||||||||||||||||
Db   290 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 349

Qy   339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
          |  |||||||||| ||||||||| ||||| ||||| |||||||| || ||||| || ||||| ||
Db   350 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 409

Qy   399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
          | || ||||| |||||||||||||||||||||||||| |||| |||||||| || ||||| |||||||||||
Db   410 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCT 469

Qy   459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
          |   ||||   | |||||   |||| | |||||||| ||||||||  | |   | || ||||
Db   470 CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 529

Qy   519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
          ||||    || || ||    | |||| | |||| |  ||||||||||||||||||| |||| |
Db   530 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 589

Qy   579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
          |||| |||||||||||| || | |   |  || || |||||| ||||| ||||||| | |||||
Db   590 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 649

Qy   639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
          |||| ||||| || ||||||||| |     ||||| ||||| |   |||| || |||
Db   650 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 709

Qy   699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
          |||| ||||| || || || || |  | |||| || ||||| || || |||||||||
Db   710 CTGCTCTGCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 769

Qy   759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
          | || || |||||||||| ||||| ||||     |||||||  |||||||| ||||| ||||
Db   770 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 829

Qy   819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
          |    ||  |   || || ||| | |||||| || | ||||  |||||||||||||
Db   830 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT 889
```

```
Qy    876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
          ||||||||||||||||||| ||||||| || ||||||| || || ||||||| || | ||
Db    890 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 949

Qy    936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
          | || ||||||||| | | | | ||||| ||||||| |||||||||||||| || | ||
Db    950 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 1009

Qy    996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
          | ||| || | |     |||| |||| |||| ||||    |||| ||||||
Db   1010 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAAGTGAGGGGCTTGTGAAACAG 1063
```

RESULT 12
BC030948

```
LOCUS        BC030948               1449 bp    mRNA    linear   PRI 12-NOV-2003
DEFINITION   Homo sapiens G protein-coupled receptor 91, mRNA (cDNA clone
             MGC:32514 IMAGE:4594810), complete cds.
ACCESSION    BC030948
VERSION      BC030948.1  GI:21410927
KEYWORDS     MGC.
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 1449)
  AUTHORS    Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
             Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
             Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
             Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
             Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
             Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
             Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
             Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
             Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
             McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
             Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
             Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
             Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
             Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
             Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
             Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
             Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
             Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE      Generation and initial analysis of more than 15,000 full-length
             human and mouse cDNA sequences
  JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE    22388257
   PUBMED    12477932
REFERENCE    2  (bases 1 to 1449)
  AUTHORS    Strausberg,R.
  TITLE      Direct Submission
  JOURNAL    Submitted (03-JUN-2002) National Institutes of Health, Mammalian
             Gene Collection (MGC), Cancer Genomics Office, National Cancer
             Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
             USA
  REMARK     NIH-MGC Project URL: http://mgc.nci.nih.gov
```

```
COMMENT         Contact: MGC help desk
                Email: cgapbs-r@mail.nih.gov
                Tissue Procurement: CLONTECH
                cDNA Library Preparation: CLONTECH Laboratories, Inc.
                cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                DNA Sequencing by: Sequencing Group at the Stanford Human Genome
                Center, Stanford University School of Medicine, Stanford, CA  94305
                Web site:       http://www-shgc.stanford.edu
                Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
                Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
                R. M.

                Clone distribution: MGC clone distribution information can be found
                through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
                Series: IRAL Plate: 41 Row: e Column: 17
                This clone was selected for full length sequencing because it
                passed the following selection criteria: matched mRNA gi: 14780893.
FEATURES             Location/Qualifiers
     source          1. .1449
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:32514 IMAGE:4594810"
                     /tissue_type="Kidney"
                     /clone_lib="NIH_MGC_75"
                     /lab_host="DH10B"
                     /note="Vector: pDNR-LIB"
     gene            1. .1449
                     /gene="GPR91"
                     /db_xref="LocusID:56670"
                     /db_xref="MIM:606381"
     CDS             112. .1104
                     /codon_start=1
                     /product="G protein-coupled receptor 91"
                     /protein_id="AAH30948.2"
                     /db_xref="GI:37572249"
                     /db_xref="LocusID:56670"
                     /translation="MAWNATCKNWLAAEAALEKYYLSIFYGIEFVVGVLGNTIVVYGY
                     IFSLKNWNSSNIYLFNLSVSDLAFLCTLPMLIRSYANGNWIYGDVLCISNRYVLHANL
                     YTSILFLTFISIDRYLIIKYPFREHLLQKKEFAILISLAIWVLVTLELLPILPLINPV
                     ITDNGTTCNDFASSGDPNYNLIYSMCLTLLGFLIPLFVMCFFYYKIALFLKQRNRQVA
                     TALPLEKPLNLVIMAVVIFSVLFTPYHVMRNVRIASRLGSWKQYQCTQVVINSFYIVT
                     RPLAFLNSVINPVFYFLLGDHFRDMLMNQLRHNFKSLTSFSRWAHELLLSFREK"
     misc_feature    217. .984
                     /note="7tm_1; Region: 7 transmembrane receptor (rhodopsin
                     family)"
                     /db_xref="CDD:pfam00001"
ORIGIN


   Query Match              38.4%;  Score 592.4;  DB 9;  Length 1449;
   Best Local Similarity    75.3%;  Pred. No. 1.6e-121;
   Matches 764;  Conservative    0;  Mismatches 246;  Indels    4;  Gaps    2;

Qy        39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
             | |   ||||||  ||||   | ||||   | || |||  |||||| ||||||||  | ||| |
Db       107 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 166
```

```
Qy    99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
         |||||||||| ||      |||||||  || |||||||  || | ||| | ||||| |||    ||
Db   167 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 226

Qy   159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
         ||| || | |||||||| ||||| ||||||      ||||||||||||||||||| ||| | |||||| |
Db   227 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 286

Qy   219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
         ||||||| ||  |||||||||||| ||||| |||||||||||| ||||| ||||||| |||||
Db   287 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 346

Qy   279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
         |||||||| || ||| ||||||| || |||||| |||||||||||||||||||||||||||||||
Db   347 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 406

Qy   339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
         | ||||||||||| ||||||||| ||||| |||||||| || ||||| || ||||| ||
Db   407 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 466

Qy   399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
         | || |||||| ||||||||||||||||||| |||| |||||||| || ||||| |||||||||||
Db   467 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCT 526

Qy   459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
         |  |||| | ||||   |||||  |||||||| | ||||||||| || |  | || ||| |||
Db   527 CCTTGGCCATTTGGGTTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 586

Qy   519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
         ||||   ||  || |||  | |||| | |||| | || | |||||||||||||||| |||| |
Db   587 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 646

Qy   579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
         |||| ||||||||||| | || || || || |||||||| ||||| |||||||| | |||||
Db   647 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 706

Qy   639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
         |||| |||||| || |||||||| |  |||||| |     ||||| ||||| |||||    |||| || |
Db   707 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 766

Qy   699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
         |||| ||||| || || || || | |     | |||| || ||||| || || ||||||||||
Db   767 CTGCTCTGCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 826

Qy   759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
         | || || |||||||||| |||||||||| ||| ||||||||| |||||||| ||||||| ||||
Db   827 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 886

Qy   819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
         | || |      || || ||| | |||||| || | |||| |||||||||||
Db   887 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT 946

Qy   876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
         ||||||||||||||||||| ||||||| || |||||||| || || ||||||| || | ||
Db   947 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTGGGAGATCACTTCA 1006

Qy   936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
```

```
       | || |||||||| | | | | ||||| ||||||| ||||||||||||||| || | ||
Db  1007 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 1066

Qy   996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
       | |||| || | | |||| |||| |||| |||| |||| ||||||
Db  1067 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAAGTGAGGGGCTTGTGAAACAG 1120
```

RESULT 13
AX342665

```
LOCUS       AX342665               1542 bp    DNA     linear   PAT 12-JAN-2002
DEFINITION  Sequence 20 from Patent WO0198351.
ACCESSION   AX342665
VERSION     AX342665.1  GI:18152045
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lal,P., Baughn,M.R., Hafalia,A.J., Nguyen,D.B., Gandhi,A.R.,
            Kallick,D.A., Griffin,J.A., Yue,H., Khan,F.A., Patterson,C.,
            Lu,D.A., Tribouley,C.M., Lu,Y., Walia,N.K., Graul,R., Yao,M.G.,
            Yang,J., Ramkumar,J., Au-Young,J., Hernandez,R., Walsh,R.T. and
            Borowsky,M.L.
  JOURNAL   Patent: WO 0198351-A 20 27-DEC-2001;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1542
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 3485895CB1"
ORIGIN
```

```
   Query Match          38.4%;  Score 592.4;  DB 6;  Length 1542;
   Best Local Similarity 75.3%;  Pred. No. 1.6e-121;
   Matches  764; Conservative   0; Mismatches  246; Indels    4; Gaps    2;

Qy    39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
          | |  |||||| |||| | |||| | || ||| ||||| ||||||| | || | |
Db   205 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 264

Qy    99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
          |||||||||| ||   ||||||| || ||||||| || ||| | ||||| ||| ||
Db   265 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 324

Qy   159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
          ||| || | ||||||| |||||| ||||||||||||||||||| ||| | ||||| |
Db   325 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 384

Qy   219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
          ||||||| || |||||||||| ||||| |||||||||||| ||||| |||||||| |||||
Db   385 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 444

Qy   279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
          |||||||||| ||  ||| |||||||| || ||||| |||||||||||||||||||||||
```

```
Db      445 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 504

Qy      339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
            |  |||||||||| |||||||| ||||| |||||||| || ||||| || ||||| ||
Db      505 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 564

Qy      399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
            | || |||||| |||||||||||||||| |||| |||||||| || |||||| ||||||||||
Db      565 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAGAAAGAGTTTGCTATTTTAATCT 624

Qy      459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
            |  |||| | ||||  ||||| ||||||||| | |||||||| || |  | || |||
Db      625 CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 684

Qy      519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
            ||||    || ||  | ||||  | ||| |  ||| |||||||||||||||| ||||  |
Db      685 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 744

Qy      579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
            ||||  |||||||||||| | || || || |||||||| ||||| |||||||| | |||||
Db      745 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 804

Qy      639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
            |||| |||||| || ||||||||| |  ||||| ||||| ||||| ||||| ||||  || |
Db      805 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 864

Qy      699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
            |||| |||||| || || || || || |    | |||| || ||||| || || ||||||||
Db      865 CTGCTCTGCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 924

Qy      759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
            | || || |||||||||||| ||||||| | ||||||||||||| ||||||||||| ||||
Db      925 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 984

Qy      819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
            | || | || || || ||| | ||||||| || | |||| ||||||||||||
Db      985 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT 1044

Qy      876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
            ||||||||||||||||||| |||||| || ||||||| || || ||||||| || | ||
Db     1045 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 1104

Qy      936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
            | || |||||||| | | | ||||| ||||||| |||||||||||||||||||| || | ||
Db     1105 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 1164

Qy      996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
            | ||| || | | |||| |||| |||| ||||    |||| ||||||
Db     1165 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAGTGAGGGGCTTGTGAAACAG 1218
```

RESULT 14
AC116026
LOCUS   AC116026   90343 bp   DNA   linear   PRI 09-APR-2002
DEFINITION Homo sapiens 3 BAC RP11-3F11 (Roswell Park Cancer Institute Human
            BAC Library) complete sequence.
ACCESSION   AC116026

```
VERSION      AC116026.1  GI:19697319
KEYWORDS     HTG.
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 90343)
  AUTHORS    Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
             Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
             Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
             Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
             Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
             Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
             Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
             Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
             Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
             Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
             Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
             Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
             Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
             Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
             Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
             Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X.,
             Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
             Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
             Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
             Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
             Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
             Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
             Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
             Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
             Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
             Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
             Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,
             Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N., Nguyen,N.,
             Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N.,
             Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
             Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M.,
             Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S.,
             Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E.,
             Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
             Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
             Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L.,
             Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S.,
             Warren,R., Washington,C., Watlington,S., Williams,G.,
             Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
             Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstock,G. and
             Gibbs,R.
  TITLE      Direct Submission
  JOURNAL    Unpublished
REFERENCE    2  (bases 1 to 90343)
  AUTHORS    Worley,K.C.
  TITLE      Direct Submission
  JOURNAL    Submitted (23-MAR-2002) Human Genome Sequencing Center, Department
             of Molecular and Human Genetics, Baylor College of Medicine, One
             Baylor Plaza, Houston, TX 77030, USA
REFERENCE    3  (bases 1 to 90343)
```

AUTHORS   Worley,K.C.
TITLE   Direct Submission
JOURNAL   Submitted (09-APR-2002) Human Genome Sequencing Center, Department
           of Molecular and Human Genetics, Baylor College of Medicine, One
           Baylor Plaza, Houston, TX 77030, USA
COMMENT   INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
            gc-help@bcm.tmc.edu

            CLONE LENGTH: This sequence does not necessarily represent the
            entire insert of this clone. Overlapping regions of clones are only
            sequenced and submitted once, so the sequence for the remainder of
            the insert may be found in the record for the adjacent clones.
            Overlapping clones are noted at the beginning and end of the
            Features listing.

            ANNOTATION OF FEATURES:
               STSs are identified using ePCR (Genome Res. 7:541-550) searches
            of a local database that includes entries from dbSTS, GDB, and
            local mapping efforts.
               Repeats are identified using RepeatMasker (A. Smit and P. Green,
            unpublished.) for Human and Mouse sequences.
               Genes and Region of sequence similarity are identified by BLAST
            (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
            EST and cDNA sequences. Genes demonstrate at least two exons
            flanked by consensus splice sites that maintained sequence
            continuity across the splice junctions. Sequences that are not
            identical matches are annotated as similar.

            SEQUENCING READ COVERAGE:Sequencing is completed to a minimum
            standard of double strand coverage with a minimum of 2 clones and 2
            reads with no ambiguities or 2 chemistries with a minimum of 2
            clones and 3 reads with no ambiguities. If the sequence quality for
            a region does not meet this standard, it will be indicated in the
            annotation as Low Coverage.

            QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality
            standards - estimated error rate less than 1 per 10,000 bases.
            Reports of lowest quality individual bases and measures of base
            quality are listed below. Description of the metrics can be found
            at URL:
            http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

            QUALSTAT-REPORT.
FEATURES   Location/Qualifiers
     source   1. .90343
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /clone="RP11-3F11"
     repeat_region   991. .1106
                     /rpt_family="MER45B"
     repeat_region   complement(1314. .1627)
                     /rpt_family="AluSx"
     repeat_region   complement(2137. .2430)
                     /rpt_family="AluY"
     repeat_region   complement(2568. .2741)

```
                             /rpt_family="L1M4"
repeat_region       complement(2742..3047)
                             /rpt_family="AluSx"
repeat_region       complement(3048..3165)
                             /rpt_family="L1M4"
repeat_region       4735..4865
                             /rpt_family="FLAM_C"
repeat_region       5657..5762
                             /rpt_family="L1MC/D"
repeat_region       5906..6237
                             /rpt_family="LTR21B"
repeat_region       6289..6773
                             /rpt_family="HERVFH21"
repeat_region       complement(8725..9597)
                             /rpt_family="MER11D"
STS                 12399..12689
                             /standard_name="136046"
repeat_region       13774..13816
                             /rpt_family="Alu"
repeat_region       13817..13874
                             /rpt_family="(TA)n"
repeat_region       complement(15157..15633)
                             /rpt_family="L2"
repeat_region       15706..15747
                             /rpt_family="AT_rich"
repeat_region       16025..16235
                             /rpt_family="MIR"
repeat_region       16560..16682
                             /rpt_family="L2"
repeat_region       complement(16710..17265)
                             /rpt_family="LTR49"
repeat_region       18077..18368
                             /rpt_family="AluSx"
repeat_region       complement(18376..18471)
                             /rpt_family="L2"
repeat_region       complement(18486..18859)
                             /rpt_family="MER57B"
repeat_region       complement(20618..20922)
                             /rpt_family="AluSx"
repeat_region       21337..21363
                             /rpt_family="AT_rich"
repeat_region       22155..22561
                             /rpt_family="L1M4"
repeat_region       complement(22608..22659)
                             /rpt_family="L1M4"
repeat_region       22685..23013
                             /rpt_family="L1MB8"
repeat_region       23103..23399
                             /rpt_family="AluSg"
repeat_region       23500..23973
                             /rpt_family="L1ME3A"
repeat_region       complement(24027..24305)
                             /rpt_family="L1MB1"
repeat_region       24304..24655
                             /rpt_family="L1ME3A"
repeat_region       24656..24697
                             /rpt_family="MADE1"
```

```
          repeat_region    25203. .25518
                           /rpt_family="AluJo"
          repeat_region    25783. .25817
                           /rpt_family="(TAA)n"
          repeat_region    26187. .26211
                           /rpt_family="AT_rich"
          repeat_region    27014. .27030
                           /rpt_family="AT_rich"
          repeat_region    complement(27031. .27316)
                           /rpt_family="AluSx"
          repeat_region    27317. .27328
                           /rpt_family="AT_rich"
          repeat_region    27574. .27615
                           /rpt_family="(TAGA)n"
          STS              28062. .28166
                           /standard_name="24707"
          STS              28199. .28382
                           /standard_name="13170"
          repeat_region    complement(29079. .29167)
                           /rpt_family="MLT1J"
          repeat_region    29168. .29532
                           /rpt_family="THE1B"
          repeat_region    complement(29533. .29552)
                           /rpt_family="MLT1J"
          repeat_region    29807. .30387
```

```
    Query Match               38.3%;  Score 590.2;  DB 9;  Length 90343;
    Best Local Similarity     75.5%;  Pred. No. 5.8e-121;
    Matches  760;  Conservative    0;  Mismatches  243;  Indels   4;  Gaps    2;

Qy        46 GGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTA 105
             ||||  |||| | |||| | || ||| ||||| |||||||| | || | ||||||||
Db     80664 GGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAAAGTACTA 80723

Qy       106 CCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGT 165
             ||| ||    ||||||| || |||||| || | ||| | ||||| ||| || ||| ||
Db     80724 CCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCATTGTTGT 80783

Qy       166 GTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCT 225
              | |||||||| ||||||    |||||||||||||||||||| ||| | ||||| ||||||||
Db     80784 TTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCTTTAACCT 80843

Qy       226 TTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAA 285
             ||  ||||||||||| |||||  ||||||||||| ||||| |||||||| ||||||||||||
Db     80844 CTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTTATGCCAA 80903

Qy       286 TGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAA 345
             ||  ||   |||  |||||||| || ||||| ||||||||||||||||||||||||  ||||
Db     80904 TGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTCATGCCAA 80963

Qy       346 CCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGCTCATGAA 405
             ||||||  |||||||| || |||| | ||||| || ||||| || ||||| || | || ||
Db     80964 CCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGATAATTAA 81023

Qy       406 GTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCTCGCTGGC 465
             ||| |||||||||||||||| |||| | |||||||| || ||||| |||||||||||  ||||
Db     81024 GTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCTCCTTGGC 81083
```

```
Qy    466 TGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATTCTGTCCC 525
              |  |||  ||||| ||||||||  | |||||||| ||  |   |  || ||| ||||
Db  81084 CATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATCCTGTTAT 81143

Qy    526 AAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAACACAATCT 585
              ||  ||  | |||| | |||  || ||||||||||||| |||| |  |||| ||
Db  81144 AACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACTACAACCT 81203

Qy    586 CATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGATGTGCTT 645
          ||||||||| | || || || ||  ||||||| ||||| ||||||||| | ||||||||| ||
Db  81204 CATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGATGTGTTT 81263

Qy    646 CTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAACTGCCCT 705
          ||| || |||||||| |   ||||| ||||| ||||| ||||      ||| || ||||| ||
Db  81264 CTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTACTGCTCT 81323

Qy    706 GCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTATACTCTT 765
          ||| || || || || |     | |||| || |||| || || |||||||||| | || ||
Db  81324 GCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTGTGCTTTT 81383

Qy    766 CACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTTG---GCC 822
          ||||||||||  ||||||||  ||||||| ||||||||  ||||||||||  ||||||   ||
Db  81384 TACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTTGGAAGCA 81443

Qy    823 ACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTCTGGCCTT 882
          |     || || |||    | ||||||| ||  | ||||         ||||||||||| |||||||
Db  81444 GTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTTTGGCCTT 81503

Qy    883 TCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACAGAGAGAT 942
          |||||||||||| |||||| ||  ||||||| || ||  ||||||| || | ||| || ||
Db  81504 TCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTGGGAGATCACTTCAGGGACAT 81563

Qy    943 GCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACATGAGCTGC 1002
          ||||||  |  |  | |||||  |||||||| |||||||||||||||||| ||  |  ||| |||
Db  81564 GCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGATGGGCTCA 81623

Qy   1003 TGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
          || |  |     |||| |||| |||| ||||  |||| |||||||
Db  81624 TGAACTCCTACTTTCATTCAGAGAAAAGTGAGGGGCTTGTGAAACAG 81670
```

RESULT 15
AC068647

```
LOCUS       AC068647              132745 bp    DNA     linear   PRI 24-JUL-2002
DEFINITION  Homo sapiens 3 BAC RP11-64D22 (Roswell Park Cancer Institute Human
            BAC Library) complete sequence.
ACCESSION   AC068647
VERSION     AC068647.10  GI:19774263
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 132745)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
```

Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,
Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N., Nguyen,N.,
Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N.,
Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M.,
Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S.,
Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E.,
Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L.,
Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S.,
Warren,R., Washington,C., Watlington,S., Williams,G.,
Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstock,G. and
Gibbs,R.
  TITLE   Direct Submission
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 132745)
  AUTHORS   Worley,K.C.
  TITLE   Direct Submission
  JOURNAL   Submitted (06-MAY-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3 (bases 1 to 132745)
  AUTHORS   Worley,K.C.
  TITLE   Direct Submission
  JOURNAL   Submitted (26-MAR-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   4 (bases 1 to 132745)
  AUTHORS   Worley,K.C.
  TITLE   Direct Submission
  JOURNAL   Submitted (28-MAR-2002) Human Genome Sequencing Center, Department

                     of Molecular and Human Genetics, Baylor College of Medicine, One
                     Baylor Plaza, Houston, TX 77030, USA
  REFERENCE   5 (bases 1 to 132745)
    AUTHORS   Worley,K.C.
    TITLE   Direct Submission
    JOURNAL   Submitted (29-MAR-2002) Human Genome Sequencing Center, Department
              of Molecular and Human Genetics, Baylor College of Medicine, One
              Baylor Plaza, Houston, TX 77030, USA
  REFERENCE   6 (bases 1 to 132745)
    AUTHORS   Worley,K.C.
    TITLE   Direct Submission
    JOURNAL   Submitted (25-JUN-2002) Human Genome Sequencing Center, Department
              of Molecular and Human Genetics, Baylor College of Medicine, One
              Baylor Plaza, Houston, TX 77030, USA
  REFERENCE   7 (bases 1 to 132745)
    AUTHORS   Worley,K.C.
    TITLE   Direct Submission
    JOURNAL   Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
              of Molecular and Human Genetics, Baylor College of Medicine, One
              Baylor Plaza, Houston, TX 77030, USA
  COMMENT   On Mar 28, 2002 this sequence version replaced gi:19718616.
              INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
              gc-help@bcm.tmc.edu

              CLONE LENGTH: This sequence does not necessarily represent the
              entire insert of this clone. Overlapping regions of clones are only
              sequenced and submitted once, so the sequence for the remainder of
              the insert may be found in the record for the adjacent clones.
              Overlapping clones are noted at the beginning and end of the
              Features listing.

              ANNOTATION OF FEATURES:
                 STSs are identified using ePCR (Genome Res. 7:541-550) searches
              of a local database that includes entries from dbSTS, GDB, and
              local mapping efforts.
                 Repeats are identified using RepeatMasker (A. Smit and P. Green,
              unpublished.) for Human and Mouse sequences.
                 Genes and Region of sequence similarity are identified by BLAST
              (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
              EST and cDNA sequences. Genes demonstrate at least two exons
              flanked by consensus splice sites that maintained sequence
              continuity across the splice junctions. Sequences that are not
              identical matches are annotated as similar.

              SEQUENCING READ COVERAGE:Sequencing is completed to a minimum
              standard of double strand coverage with a minimum of 2 clones and 2
              reads with no ambiguities or 2 chemistries with a minimum of 2
              clones and 3 reads with no ambiguities. If the sequence quality for
              a region does not meet this standard, it will be indicated in the
              annotation as Low Coverage.

              QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality
              standards - estimated error rate less than 1 per 10,000 bases.
              Reports of lowest quality individual bases and measures of base
              quality are listed below. Description of the metrics can be found
              at URL:
              http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

```
            QUALSTAT-REPORT.
FEATURES                 Location/Qualifiers
    source               1. .132745
                         /organism="Homo sapiens"
                         /mol_type="genomic DNA"
                         /db_xref="taxon:9606"
                         /chromosome="3"
                         /clone="RP11-64D22"
    misc_feature         1. .2005
                         /note="overlaps bases 170209. .172213 of clone AC069067"
                         /function="clone overlap"
    STS                  30. .130
                         /standard_name="74493"
    repeat_region        complement(522. .1015)
                         /rpt_family="MLT1D"
    repeat_region        complement(2452. .2697)
                         /rpt_family="L1MA5A"
    repeat_region        complement(3200. .3578)
                         /rpt_family="MLT1B"
    repeat_region        3600. .3749
                         /rpt_family="(TA)n"
    repeat_region        4391. .4411
                         /rpt_family="AT_rich"
    repeat_region        4909. .4960
                         /rpt_family="AT_rich"
    repeat_region        complement(5657. .6403)
                         /rpt_family="L1PA13"
    repeat_region        6404. .7799
                         /rpt_family="L1PA13"
    STS                  8045. .8318
                         /standard_name="183647"
    repeat_region        complement(8708. .9282)
                         /rpt_family="L1MD3"
    repeat_region        complement(9287. .9357)
                         /rpt_family="MLT1F1"
    repeat_region        complement(9359. .9460)
                         /rpt_family="L1MC3"
    repeat_region        complement(9587. .9880)
                         /rpt_family="AluSg"
    repeat_region        complement(10203. .10450)
                         /rpt_family="L1MC4"
    repeat_region        11468. .11699
                         /rpt_family="L1M4"
    repeat_region        11717. .11886
                         /rpt_family="MLT2CB"
    repeat_region        11908. .11982
                         /rpt_family="Tigger3(Golem)"
    repeat_region        12020. .12246
                         /rpt_family="THE1C"
    repeat_region        12263. .12562
                         /rpt_family="AluSx"
    repeat_region        13326. .13346
                         /rpt_family="AT_rich"
    repeat_region        complement(13934. .14245)
                         /rpt_family="AluSx"
    repeat_region        complement(14256. .14568)
```

```
                              /rpt_family="AluJo"
         repeat_region       14618. .14746
                              /rpt_family="MER8"
         repeat_region       14825. .14849
                              /rpt_family="AT_rich"
         repeat_region       14865. .14906
                              /rpt_family="AT_rich"
         repeat_region       15465. .15739
                              /rpt_family="L2"
         repeat_region       16579. .16756
                              /rpt_family="(TTATA)n"
         repeat_region       complement(16757. .17074)
                              /rpt_family="L2"
         repeat_region       17621. .17660
                              /rpt_family="(CAAAA)n"
         repeat_region       18544. .18725
```

```
   Query Match                38.3%;  Score 590.2;  DB 9;   Length 132745;
   Best Local Similarity      75.5%;  Pred. No. 5.9e-121;
   Matches  760;  Conservative   0;  Mismatches  243;  Indels   4;  Gaps   2;
```

```
Qy        46 GGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTA 105
             ||||  ||||  | ||||  | ||  ||| |||||  |||||||||  | || | |||||||||
Db    123065 GGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAAAGTACTA
      123124

Qy       106 CCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGT 165
             ||| ||      ||||||||  ||  ||||||| || | |||  | |||||| |||  || ||
Db    123125 CCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCATTGTTGT
      123184

Qy       166 GTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCT 225
             | |||||||| ||||||      ||||||||||||||||||| ||| | |||||| ||||||||
Db    123185 TTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCTTTAACCT
      123244

Qy       226 TTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAA 285
             ||   |||||||||| |||||  |||||||||||| ||||| |||||||| |||||||||||||
Db    123245 CTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTTATGCCAA
      123304

Qy       286 TGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAA 345
             ||   ||   |||  ||||||||| || |||||  ||||||||||||||||||||||  ||||
Db    123305 TGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTCATGCCAA
      123364

Qy       346 CCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGCTCATGAA 405
             ||||||| |||||||| |||||  ||||| |||||| || ||||||  || | || ||
Db    123365 CCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGATAATTAA
      123424

Qy       406 GTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCTCGCTGGC 465
             ||| ||||||||||||||||| |||| ||||||| ||||||| ||||||||||||||  ||||
Db    123425 GTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCTCCTTGGC
      123484

Qy       466 TGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATTCTGTCCC 525
```

```
                     |  ||||   |||||  ||||||||   |  ||||||||  ||   |   |  ||  |||  ||||
Db    123485 CATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATCCTGTTAT
123544

Qy       526 AAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAACACAATCT 585
                  ||  ||  |  ||||   |  |||  |  ||  | ||||||||||||||  ||||  |  ||||  ||
Db    123545 AACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACTACAACCT
123604

Qy       586 CATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGATGTGCTT 645
                 ||||||||||  |  ||  ||  ||    |||||||  |||||  ||||||||  |  |||||||||  ||
Db    123605 CATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGATGTGTTT
123664

Qy       646 CTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAACTGCCCT 705
                 |||  ||  |||||||||  |     |||||  |||||     |||||       ||||     ||  |||||  ||
Db    123665 CTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTACTGCTCT
123724

Qy       706 GCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTATACTCTT 765
                 |||  ||  ||  ||  ||  |  |     |  ||||  ||  ||||  ||  ||  ||||||||||  |  ||  ||
Db    123725 GCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTGTGCTTTT
123784

Qy       766 CACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTTG---GCC 822
                 ||||||||||||   |||||||  |||   ||||||||||   ||||||||||     |||||       ||
Db    123785 TACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTTGGAAGCA
123844

Qy       823 ACAAGGATGTACACAGAAGGCCATCAAATCTATATACACTGACACGGCCTCTGGCCTT 882
                 |       ||  ||  |||    |  ||||||  ||   |  ||||       ||||||||||||  |||||||
Db    123845 GTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTTTGGCCTT
123904

Qy       883 TCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACAGAGAGAT 942
                 ||||||||||||  ||||||  ||   |||||||  ||  ||   |||||||  ||  |  |||  ||  ||
Db    123905 TCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCAGGGACAT
123964

Qy       943 GCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACATGAGCTGC 1002
                 ||||||  |  |  |   |  |||||   |||||||  ||||||||||||||||||||  ||  |  |||  |||
Db    123965 GCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGATGGGCTCA
124024

Qy      1003 TGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
                 ||  |   |       ||||  ||||  ||||  ||||     ||||  ||||||
Db    124025 TGAACTCCTACTTTCATTCAGAGAAAAGTGAGGGGCTTGTGAAACAG 124071


Search completed: August 24, 2004, 14:51:12
Job time : 6260 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 08:41:34 ; Search time 655 Seconds
                                               (without alignments)
                                               10007.587 Million cell updates/sec

Title: US-09-891-138A-1
Perfect score: 1543
Sequence: 1 gctcctggcagagttttctg..........tgcctaaataaatcaatata 1543

Scoring table: IDENTITY_NUC
                   Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   N_Geneseq_29Jan04:*
                   1: geneseqn1980s:*
                   2: geneseqn1990s:*
                   3: geneseqn2000s:*
                   4: geneseqn2001as:*
                   5: geneseqn2001bs:*
                   6: geneseqn2002s:*
                   7: geneseqn2003as:*
                   8: geneseqn2003bs:*
                   9: geneseqn2003cs:*
                   10: geneseqn2004s:*

        Pred. No. is the number of results predicted by chance to have a
        score greater than or equal to the score of the result being printed,
        and is derived by analysis of the total score distribution.

                                        SUMMARIES
                        %
Result   Query
   No.   Score   Match Length DB   ID   Description
----
      1   1543   100.0   1543   6   ABK12957   Abk12957 DNA seque
      2   592.4   38.4   1005   3   AAA46036   Aaa46036 Human G p
      3   592.4   38.4   1005   3   AAD01135   Aad01135 Human orp
      4   592.4   38.4   1005   7   ACA93273   Aca93273 Human cDN
      5   592.4   38.4   1380   7   ABZ42542   Abz42542 Human pur
      6   592.4   38.4   1436   6   ABL90790   Abl90790 Human pol
      7   592.4   38.4   1473   7   ACC46165   Acc46165 Human dit

	8	592.4	38.4	1542	6	AAD24958	Aad24958 Human G-p
	9	590.8	38.3	1338	7	ABS57291	Abs57291 cDNA enco
	10	590.8	38.3	1428	8	ACD27619	Acd27619 Human ATP
	11	589.2	38.2	1996	2	AAT71900	Aat71900 Human pur
	12	587.2	38.1	1428	2	AAT75146	Aat75146 Human ATP
	13	580.4	37.6	1385	3	AAC81122	Aac81122 Human sec
	14	572.8	37.1	1005	9	ADC12679	Adc12679 Human GPC
c	15	158.8	10.3	639	9	ADE85578	Ade85578 Farnesyl
	16	126.6	8.2	1011	6	AAL43942	Aal43942 Human G p
	17	126.6	8.2	1014	5	AAS07948	Aas07948 Human cDN
	18	126.6	8.2	1014	6	ABN85630	Abn85630 Human P2Y
	19	126.6	8.2	1014	6	AAK98323	Aak98323 Human pur
	20	126.6	8.2	1014	6	ABK11381	Abk11381 Human DNA
	21	126.6	8.2	1014	6	ABZ42876	Abz42876 Human GPC
	22	126.6	8.2	1014	7	ABZ59170	Abz59170 Human TGR
	23	126.6	8.2	1014	7	ABZ42582	Abz42582 Human G p
	24	126.6	8.2	1014	9	ADC25997	Adc25997 Human pur
	25	126.6	8.2	1014	9	AAD61652	Aad61652 Human GPC
	26	126.6	8.2	1029	7	AAD50885	Aad50885 TARZAN DN
	27	126.6	8.2	1081	6	ABS59232	Abs59232 Human G-p
	28	126.6	8.2	1179	7	AAD50883	Aad50883 Human TAR
	29	126.6	8.2	1288	6	ABL56197	Abl56197 Human P2Y
	30	126.6	8.2	1414	9	ADC86010	Adc86010 Human GPC
	31	126.6	8.2	1729	5	AAS08362	Aas08362 Human cDN
	32	126.6	8.2	1729	5	ABV25767	Abv25767 Human pro
	33	126.6	8.2	1729	5	ABV30024	Abv30024 Human pro
	34	126.6	8.2	1729	5	ABV24026	Abv24026 Human pro
	35	126.6	8.2	1729	5	ABV29909	Abv29909 Human pro
	36	126.6	8.2	1797	7	AAD50882	Aad50882 Human TAR
	37	126.6	8.2	5435	6	ABL56198	Abl56198 Human P2Y
	38	126.6	8.2	9905	6	AAK98324	Aak98324 Human pur
	39	125.6	8.1	990	9	ADC12673	Adc12673 Human GPC
	40	125	8.1	1014	6	AAD26370	Aad26370 Human G-p
	41	125	8.1	1014	6	AAD34278	Aad34278 Human AXO
	42	125	8.1	1014	6	ABQ78847	Abq78847 Human G-p
	43	125	8.1	1560	6	ABS51730	Abs51730 Human nov
	44	125	8.1	1851	6	ABS51678	Abs51678 DNA encod
	45	119.2	7.7	1020	4	AAH51011	Aah51011 Human nGP

ALIGNMENTS

RESULT 1
ABK12957
ID   ABK12957 standard; DNA; 1543 BP.
XX
AC   ABK12957;
XX
DT   09-APR-2002 (first entry)
XX
DE   DNA sequence of mouse G-protein coupled receptor TGR18 gene.
XX
KW   Mouse; G-protein coupled; receptor; GPCR; TGR18; kidney disease;
KW   signal transduction modulator; cerebral cavernous malformation;
KW   hyperlipidemia; obesity; dyslexia; cardiac myxoma; renal failure;
KW   nephritis; hypertension; liver disease; cirrhosis; blood disorder;

```
Qy    1 GCTCCTGGCAGAGTTTTCTGTCGAGACAGAAGCCGACAGCAGAATGGCACAGAATTTATC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 GCTCCTGGCAGAGTTTTCTGTCGAGACAGAAGCCGACAGCAGAATGGCACAGAATTTATC 60

Qy   61 TTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTA 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTA 120

Qy  121 TGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGTGTTCGGCTACCTCTT 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 TGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGTGTTCGGCTACCTCTT 180

Qy  181 CTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCTTTCCATCTCTGACTT 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCTTTCCATCTCTGACTT 240

Qy  241 TGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAATGATAAGGGGACCTA 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 TGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAATGATAAGGGGACCTA 300

Qy  301 TGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAACCTCTACACCAGCAT 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 TGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAACCTCTACACCAGCAT 360

Qy  361 CCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGCTCATGAAGTACCCTTTCCGAGA 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 CCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGCTCATGAAGTACCCTTTCCGAGA 420

Qy  421 ACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCTCGCTGGCTGTCTGGGCCTTAGT 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 ACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCTCGCTGGCTGTCTGGGCCTTAGT 480

Qy  481 GACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATTCTGTCCCAAAAGAAGAGGGCAG 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 GACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATTCTGTCCCAAAAGAAGAGGGCAG 540

Qy  541 TAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAACACAATCTCATTTACAGCCTCTG 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 TAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAACACAATCTCATTTACAGCCTCTG 600

Qy  601 CCTGACTTTGTTGGGCTTCCTAATTCCTCTCTGTGATGTGCTTCTTCTACTACAAGAT 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 CCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGATGTGCTTCTTCTACTACAAGAT 660

Qy  661 GGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAACTGCCCTGCCACTGGACAAACC 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 GGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAACTGCCCTGCCACTGGACAAACC 720

Qy  721 CCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTATACTCTTCACACCCTATCATAT 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 CCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTATACTCTTCACACCCTATCATAT 780

Qy  781 CATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTTGGCCACAAGGATGTACACAGAA 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 CATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTTGGCCACAAGGATGTACACAGAA 840
```

```
Qy   841 GGCCATCAAATCTATATACACACTGACACGGCCTCTGGCCTTTCTGAACAGTGCCATCAA 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GGCCATCAAATCTATATACACACTGACACGGCCTCTGGCCTTTCTGAACAGTGCCATCAA 900

Qy   901 TCCCATCTTCTACTTCCTCATGGGAGACCATTACAGAGAGATGCTGATTAGTAAGTTCAG 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCCCATCTTCTACTTCCTCATGGGAGACCATTACAGAGAGATGCTGATTAGTAAGTTCAG 960

Qy   961 ACAATACTTCAAGTCCCTTACATCCTTCAGGACATGAGCTGCTGGATGCAGGTCTTCACT 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 ACAATACTTCAAGTCCCTTACATCCTTCAGGACATGAGCTGCTGGATGCAGGTCTTCACT 1020

Qy  1021 CAGCCAAAATGAGACACTTGATAAACAGTGCTGTGCAGTTGAGTTTTAACTAAGTAAACC 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 CAGCCAAAATGAGACACTTGATAAACAGTGCTGTGCAGTTGAGTTTTAACTAAGTAAACC 1080

Qy  1081 ACCATTTCTAGGCTTTAGCTTTCCACCATCCTCCAACCCCCAGGGCTGGAGTACAAGCTG 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 ACCATTTCTAGGCTTTAGCTTTCCACCATCCTCCAACCCCCAGGGCTGGAGTACAAGCTG 1140

Qy  1141 GGTCCACATGAATCAGAAGGCAGCTCTCTGTTCTGATTTTAGGTTATACCCAGAGTATGG 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 GGTCCACATGAATCAGAAGGCAGCTCTCTGTTCTGATTTTAGGTTATACCCAGAGTATGG 1200

Qy  1201 AAAAAATAAGGCATGAGAAAGCATTGACATCTTCACTTAAGAACTGAACAAAAGAGAACA 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 AAAAAATAAGGCATGAGAAAGCATTGACATCTTCACTTAAGAACTGAACAAAAGAGAACA 1260

Qy  1261 AATATTGTCAATGTTTGGACACTTAGGATCTGAAATCTTGGAAATTTTAAGACCTCTTTT 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 AATATTGTCAATGTTTGGACACTTAGGATCTGAAATCTTGGAAATTTTAAGACCTCTTTT 1320

Qy  1321 TCTATCAGTGTAAAAGGAATACAAGATAGCTAGTTGCAAATGCTGAATGCATTTCATCAT 1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 TCTATCAGTGTAAAAGGAATACAAGATAGCTAGTTGCAAATGCTGAATGCATTTCATCAT 1380

Qy  1381 TGGTCAGGTCGATAAGCGTGTTTCTGAAATAGTCTTATTTTTATTCTTGTAATATTAAAA 1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 TGGTCAGGTCGATAAGCGTGTTTCTGAAATAGTCTTATTTTTATTCTTGTAATATTAAAA 1440

Qy  1441 TTTATGTGAAAAATGAATATAATTCAATGTACAACATTAGATTTTCTATTTGAAAATTAT 1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 TTTATGTGAAAAATGAATATAATTCAATGTACAACATTAGATTTTCTATTTGAAAATTAT 1500

Qy  1501 ATTTCTTGAAAAAATAACTGCTGTGCCTAAATAAATCAATATA 1543
          |||||||||||||||||||||||||||||||||||||||||||
Db  1501 ATTTCTTGAAAAAATAACTGCTGTGCCTAAATAAATCAATATA 1543
```

RESULT 2
AAA46036
ID   AAA46036 standard; cDNA; 1005 BP.
XX
AC   AAA46036;
XX
DT   22-AUG-2000 (first entry)
XX

```
     Query Match              38.4%;   Score 592.4;   DB 3;   Length 1005;
     Best Local Similarity    75.5%;   Pred. No. 1.3e-139;
     Matches   750;  Conservative     0;  Mismatches  241;  Indels   3;  Gaps   1;

Qy        39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
              | |  ||||||  ||||  |  ||||  |  || |||  |||||  ||||||||  | ||  |  |
Db         8 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 67

Qy        99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
              ||||||||||| ||    |||||||  || |||||| || | ||| | ||||| ||| || ||
Db        68 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 127

Qy       159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
              ||| || | ||||||| ||||||     |||||||||||||||||||||||| ||| | |||||| |
Db       128 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 187

Qy       219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
              ||||||| ||  |||||||||||| ||||| ||||||||||| ||||| ||||||||| |||||
Db       188 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 247

Qy       279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
              ||||||||| || ||| ||||||||| || ||||| ||||||||||||||||||||||||||||
Db       248 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 307

Qy       339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
              |  |||||||||| |||||||||| ||||| ||||||||| || |||||| || |||||  ||
Db       308 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 367

Qy       399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
              | || |||||| ||||||||||||||||| |||| |||||||| || ||||| ||||||||||
Db       368 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCT 427

Qy       459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
              | |||| | ||||| ||||| |||||||| | ||||||||| || |  | || |||  ||
Db       428 CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 487

Qy       519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
              ||||    || ||  | ||||  | ||| |   || | ||||||||||||| |||| |
Db       488 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 547

Qy       579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
              |||| ||||||||||||| | || || || || ||||||| ||||| ||||||| | |||||
```

```
Db    548 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 607

Qy    639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
          |||| |||||| || |||||||||| |    ||||| |||||| ||||| | |||| || |
Db    608 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 667

Qy    699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
          |||| |||||| || || || || |   | |||| || |||| || || ||||||||||
Db    668 CTGCTCTGCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 727

Qy    759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
          | || || |||||||||||| ||||||| ||| ||||||||||| ||||||||| ||||
Db    728 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 787

Qy    819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
          |   || |   || || ||| | |||||| || | |||     |||||||||||
Db    788 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT 847

Qy    876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
          ||||||||||||||||||| |||||| || ||||||| || || ||||||| || | ||
Db    848 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 907

Qy    936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
          | || |||||||| | | | ||||| ||||||| ||||||||||||||||| || | ||
Db    908 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 967

Qy    996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA 1029
          | ||| || | |    |||| |||| ||||
Db    968 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAA 1001
```

RESULT 3
AAD01135
```
ID   AAD01135 standard; cDNA; 1005 BP.
XX
AC   AAD01135;
XX
DT   02-NOV-2000   (first entry)
XX
DE   Human orphan G protein-coupled receptor hCHN10 cDNA.
XX
KW   Human; orphan G protein-coupled receptor; GPCR; hCHN10; drug screening;
KW   transmembrane receptor; expressed sequence tag; EST; signal cascade; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1005
FT                   /*tag=  a
FT                   /product= "hCHN10"
FT                   /note= "Human orphan G protein-coupled receptor"
XX
PN   WO200031258-A2.
XX
PD   02-JUN-2000.
XX
PF   13-OCT-1999;   99WO-US023687.
```

CC   The present sequence is a cDNA encoding hCHN10, an endogenous human
CC   orphan G protein-coupled receptor (GPCR), expressed in kidney and
CC   thyroid. The hCHN10 cDNA was identified using the human EST (expressed
CC   sequence tag) 1365839 as a probe. The orphan GPCR of the invention, like
CC   all GPCRs has seven transmembrane alpha helices with an extracellular N-
CC   terminus and an intracellular C-terminus. However, no endogenous ligands
CC   has yet been identified for the proteins of the invention. The orphan
CC   GPCRs may be used in the identification of their endogenous ligands, and
CC   to screen potential GPCR agonists and antagonists for use as
CC   pharmaceutical agents. The proteins may also be used in the study of GPCR
CC   -mediated signalling cascades, and to elucidate their precise role in
CC   normal and diseased human conditions. Nucleic acid encoding human orphan
CC   GPCRs may be used for tissue localisation expression analysis to provide
CC   information about their function in healthy and pathological states
XX
SQ   Sequence 1005 BP; 248 A; 236 C; 196 G; 325 T; 0 U; 0 Other;

     Query Match   38.4%;   Score 592.4; DB 3; Length 1005;
     Best Local Similarity   75.5%;   Pred. No. 1.3e-139;
      Matches 750; Conservative   0; Mismatches 241; Indels   3; Gaps   1;

```
Qy    39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
         | |  |||||||   ||||  |  ||||  | || |||  |||||  ||||||||  |  || |||
Db     8 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 67

Qy    99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
         |||||||||| ||     ||||||| || |||||| || |  ||| | ||||| ||| ||
Db    68 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 127

Qy   159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
         ||| || | |||||||| ||||| ||||||      ||||||||||||||||||| |  | |||||| |
Db   128 TTGTTGTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 187

Qy   219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
         |||||||  ||   ||||||||| |||||  ||||||||||  |||||  ||||||||  |||||
Db   188 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 247

Qy   279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
         |||||||||  ||   |||   ||||||| || |||||  ||||||||||||||||||||||||
Db   248 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 307

Qy   339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
         |  |||||||||| ||||||||| ||||| ||||||||| || | ||||| || ||||||  ||
Db   308 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 367

Qy   399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
         |  ||  ||||| |||||||||||||| ||||  |||| ||||||||| || ||||| |||||||||||
Db   368 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCT 427

Qy   459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
         |  ||||  |  |||| ||||| |||| ||||||||| |  ||||||||| ||  |  | || |||
Db   428 CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 487

Qy   519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
         ||||     ||  ||  | ||||  |||| |  || | ||||||||||||||||| |||| |
Db   488 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 547

Qy   579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
         |||| |||||||||||| | ||  || || || || ||||||| ||||| |||||||| | |||||
Db   548 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 607

Qy   639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
         ||||| ||||| || |||||||| |     |||||  |||| ||||  |||||    |||| || |
Db   608 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 667

Qy   699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
         |||| ||||| || || || || |||    | |||| || || || ||||||||||||
Db   668 CTGCTCTGCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 727

Qy   759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
         | || || |||||||||| |||||| ||| ||||||||||| ||||||||||||||||||| ||||
Db   728 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 787

Qy   819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
         |  |  ||   |    || || |||   | ||||||| ||  |  |||| |||||||||||
Db   788 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT 847
```

```
Qy    876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
          |||||||||||||||||| |||||| || |||||||| || || ||||||| || | ||
Db    848 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 907

Qy    936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
          | || |||||||| | | | | |||||| ||||||| |||||||||||||| || | ||
Db    908 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 967

Qy    996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA 1029
          | ||| || | |    |||| |||| ||||
Db    968 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAA 1001
```

RESULT 4
ACA93273

```
  Query Match            38.4%;  Score 592.4;  DB 7;  Length 1005;
  Best Local Similarity  75.5%;  Pred. No. 1.3e-139;
  Matches  750;  Conservative   0;  Mismatches 241;  Indels   3;  Gaps    1;


Qy      39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
           | |  ||||||| |||| | |||| | || ||| ||||| |||||||| | || | |
Db       8 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 67


Qy      99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
           |||||||||| ||     ||||||| || |||||| || | ||| | |||||| ||| ||
Db      68 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 127


Qy     159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
           ||| || | ||||||| |||||| |||||         |||||||||||||| || | |||||| |
Db     128 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 187


Qy     219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
           ||||||| || ||||||||||| ||||| ||||||||||| ||||| |||||||| |||||
Db     188 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 247


Qy     279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
           ||||||||| || ||| ||||||||| || ||||| |||||||||||||||||||||||||
Db     248 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 307


Qy     339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
           |  |||||||||| |||||||| ||||| ||||||||| || ||||| || ||||| ||
Db     308 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 367
```

```
Qy    399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
          | || ||||| |||||||||||||||| |||| |||||||| || ||||| |||||||||||
Db    368 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCT 427

Qy    459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
          |  |||| | |||| ||||| ||||||||| | ||||||||| ||  |  | || |||
Db    428 CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 487

Qy    519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
          ||||    || || | |||| | ||| |  || | ||||||||||||||||| |||| |
Db    488 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 547

Qy    579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
          ||||  |||||||||| | || | ||  || ||   ||||||| ||||| ||||||||| | |||||
Db    548 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 607

Qy    639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
          |||| ||||| || ||||||||| |     ||||| ||||| |||||     |||| || |
Db    608 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 667

Qy    699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
          |||| ||||| || || || ||| |    | |||| || |||| || || || ||||||||||
Db    668 CTGCTCTGCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 727

Qy    759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
          | || || ||||||||||||| |||||| ||||||| ||| ||||||||||| ||||||||| ||||
Db    728 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 787

Qy    819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
          |  || | || || ||| |||  | |||||| || | ||||     |||||||||||
Db    788 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT 847

Qy    876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
          ||||||||||||||||| |||||| || |||||||| || || ||||||| || | ||
Db    848 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 907

Qy    936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
          | || |||||||||| | | |  | ||||| ||||||| |||||||||||||||| || | ||
Db    908 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 967

Qy    996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA 1029
          | |||  || |  |     |||| |||| ||||
Db    968 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAA 1001
```

RESULT 5
ABZ42542
ID   ABZ42542 standard; DNA; 1380 BP.
XX
AC   ABZ42542;
XX
DT   04-MAR-2003   (first entry)
XX
DE   Human purinergic receptor P2U2 nucleotide SEQ ID NO:566.
XX
KW   G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW   G protein-coupled receptor modulator; antibody; immune-related disease;

CC   The present invention describes antigenic peptides (I) comprising: (a)
CC   any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
CC   acids. Also described: (1) an assay for the detection of a particular G
CC   protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC   and (2) an isolated antibody having high specificity and high affinity or
CC   avidity for a particular GPCR. (I) can be used as GPCR modulators and in
CC   gene therapy. The antigenic peptides for GPCRs are useful in detecting an
CC   antibody against a particular GPCR, and in the production of specific
CC   antibodies. The peptides and antibodies are also useful for detecting the
CC   presence or absence of corresponding GPCRs. The antigenic peptides for
CC   GPCRs and antibodies are useful for diagnosing and designing drugs for
CC   treating immune-related diseases, growth-related diseases, cell
CC   regeneration-related disease, immunological-related cell proliferative
CC   diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC   atherosclerosis, bacterial, fungal, protozoan or viral infections,
CC   osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC   inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC   disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC   anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC   loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC   hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC   any other disorder in which GPCRs are involved. The antibodies may be
CC   used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode

SQ   Sequence 1380 BP; 383 A; 294 C; 274 G; 429 T; 0 U; 0 Other;

      Query Match   38.4%; Score 592.4; DB 7; Length 1380;
      Best Local Similarity 75.3%; Pred. No. 1.5e-139;
      Matches 764; Conservative   0; Mismatches 246; Indels   4; Gaps   2;

Qy   39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
              | | |||||| |||| | |||| | || ||| ||||| |||||||| | || | |
Db   50 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 109

Qy   99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
             |||||||||| || ||||||| || |||||| || | ||| | ||||| ||| ||
Db   110 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 169

Qy   159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
             ||| || | ||||||| |||||| |||||||||||||||||||| ||| | ||||| |
Db   170 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 229

Qy   219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
             |||||| || ||||||||||| ||||| ||||||||||| ||||| ||||||| |||||
Db   230 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 289

Qy   279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
             ||||||||| || ||| |||||||| || ||||| |||||||||||||||||||||||||
Db   290 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 349

Qy   339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
             | |||||||||| |||||||||| |||||| ||||||||| ||||| || ||||| || ||
Db   350 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 409

Qy   399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
             | || |||||| |||||||||||||||||| |||| ||||||||| || |||||| ||||||||||||
Db   410 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAGAAAGAGTTTGCTATTTTAATCT 469

Qy   459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
             | |||| | |||| ||||| |||||||| | ||||||||| || | | || |||
Db   470 CCTTGGCCATTTGGGTTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 529

Qy   519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
             |||| || || | |||| | ||| | || | |||||||||||||| |||| |
Db   530 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 589

Qy   579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
             |||| |||||||||||| || | | | || ||||||| ||||| |||||||| | ||||||
Db   590 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 649

Qy   639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
             |||| ||||| || ||||||||| | ||||| |||||| || ||| | |||| | |
Db   650 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 709

Qy   699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
             |||| |||||| || || || || | | | | |||| || |||| | | |||||||||
Db   710 CTGCTCTGCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 769

```
Qy    759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
           | || || |||||||||||  ||||||| ||| |||||||||| ||||||||||  ||||
Db    770 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 829

Qy    819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
           |   ||   |    || || |||  | |||||| ||  | |||    |||||||||||
Db    830 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT 889

Qy    876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
           |||||||||||||||||| ||||||| || ||||||| || || ||||||| || | ||
Db    890 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 949

Qy    936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
           | || ||||||||| | | | ||||| ||||||| |||||||||||||||| || | ||
Db    950 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 1009

Qy    996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
           | ||| || | |     |||| |||| |||| |||| |||    ||||  ||||||
Db   1010 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAGTGAGGGGCTTGTGAAACAG 1063
```

RESULT 6
ABL90790

```

```
PT   disorders.
XX
PS   Claim 4; SEQ ID NO 1352; 2081pp + Sequence Listing; English.
XX
CC   The invention relates to novel genes (ABL89449-ABL90853) and proteins
CC   (ABB89040-ABB90444) useful for preventing, treating or ameliorating
CC   medical conditions e.g. by protein or gene therapy. The genes are
CC   isolated from a range of human tissues disclosed in the specification.
CC   The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC   the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC   ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC   breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC   disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC   anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC   multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC   cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC   ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC   infectious diseases such as viral, bacterial, fungal and parasitic
CC   infections. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1436 BP; 397 A; 309 C; 289 G; 441 T; 0 U; 0 Other;

      Query Match   38.4%;   Score 592.4;   DB 6;   Length 1436;
      Best Local Similarity 75.3%;   Pred. No. 1.6e-139;
       Matches 764; Conservative   0; Mismatches 246; Indels   4; Gaps   2;

Qy   39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
            | | ||||||| |||| | |||| | || ||| ||||| |||||||| | ||| |
Db   100 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 159

Qy   99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
            |||||||||| || ||||||| || |||||| || | ||| | |||||| ||| ||
Db   160 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 219

Qy   159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
            ||| || | ||||||| |||||| |||||||||||||||||||| ||| | ||||| |
Db   220 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 279

Qy   219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
            ||||||| || ||||||||| ||||| ||||||||||| ||||| ||||||| |||||
Db   280 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 339

Qy   279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
            |||||||| || ||| |||||||| || ||||| |||||||||||||||||||||||||||
Db   340 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 399

Qy   339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
            | ||||||||| ||||||| ||||| ||||| ||||||| || ||||| || ||||| ||
Db   400 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 459

Qy   399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
            | || ||||| |||||||||||||||||| |||| ||||||| || ||||| ||||||||||
Db   460 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAGAAAGAGTTTGCTATTTTAATCT 519

Qy   459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
```

```
           |   ||||  |  ||||   |||||  ||||||||  |  ||||||||  ||   |   |  ||  |||
Db     520 CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 579

Qy     519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
           ||||   ||  ||  |  |||||  |  |||  |   ||  |  ||||||||||||||  ||||  |
Db     580 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 639

Qy     579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
           ||||  |||||||||||||  |  ||  ||  ||   |||||||  |||||  ||||||||  |  |||||
Db     640 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 699

Qy     639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
           ||||  ||||||  ||  |||||||||  |   |||||  |||||  ||||||  ||||  |  |
Db     700 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 759

Qy     699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
           ||||  |||||  ||  ||  ||  ||  |   |  ||||   ||  ||||  ||  ||  ||||||||||
Db     760 CTGCTCTGCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 819

Qy     759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
           |  ||  ||  |||||||||||  |||||||  |||  ||||||||||  |||||||||  ||||
Db     820 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 879

Qy     819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
           |   ||  |     ||  ||  |||   |  ||||||  ||  |  ||||      ||||||||||||
Db     880 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT 939

Qy     876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
           |||||||||||||||||||  |||||  ||   ||||||  ||  ||   |||||||  ||  |  ||
Db     940 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 999

Qy     936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
           |  ||  |||||||||  |  |   |  |||||  |||||||  ||||||||||||||||  ||  |  ||
Db    1000 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 1059

Qy     996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
           |  |||  ||  |   |      ||||  ||||   ||||  ||||      ||||  ||||||
Db    1060 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAAGTGAGGGGCTTGTGAAACAG 1113
```

RESULT 7
ACC46165
ID   ACC46165 standard; cDNA; 1473 BP.
XX
AC   ACC46165;
XX
DT   02-JUN-2003   (first entry)
XX
DE   Human dithp receptor-encoding cDNA.
XX
KW   Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;
KW   cancer; cell proliferative disorder; autoimmune disorder;
KW   inflammatory disorder; infection; hormonal disorder; metabolic disorder;
KW   neurological disorder; gastrointestinal disorder; transport disorder;
KW   connective tissue disorder; drug screening; proteome analysis;
KW   gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
KW   disease model; toxicological testing; transcript imaging; receptor; gene;

```
KW    ss.
XX
OS    Homo sapiens.
XX
PN    WO200297031-A2.
XX
PD    05-DEC-2002.
XX
PF    27-MAR-2002; 2002WO-US010056.
XX
PR    28-MAR-2001; 2001US-0279619P.
PR    29-MAR-2001; 2001US-0280067P.
PR    29-MAR-2001; 2001US-0280068P.
PR    16-MAY-2001; 2001US-0291280P.
PR    17-MAY-2001; 2001US-0291829P.
PR    17-MAY-2001; 2001US-0291849P.
PR    19-JUN-2001; 2001US-0299428P.
PR    20-JUN-2001; 2001US-0299776P.
PR    20-JUN-2001; 2001US-0300001P.
XX
PA    (INCY-) INCYTE GENOMICS INC.
XX
PI    Daffo A,  Jones AL,  Tran AB,  Dahl CR,  Gietzen D,  Chinn J;
PI    Dufour GE,  Hillman JL,  Yu JY,  Tuason O,  Yap PE,  Amshey SR;
PI    Daughtery SC,  Dam TC,  Liu TF,  Nguyen DA,  Kleefeld Y,  Gerstin EH;
PI    Peralta CH,  David MH,  Lewis SA,  Chen AJ,  Panzer SR,  Harris B;
PI    Flores V,  Marwaha R,  Lo A,  Lan RY,  Urashka ME;
XX
DR    WPI; 2003-129518/12.
DR    P-PSDB; ABR41222.
XX
PT    Novel human diagnostic and therapeutic polypeptide useful for identifying
PT    test compound which specifically binds to a polypeptide encoded by human
PT    diagnostic and therapeutic polynucleotide, and to induce antibodies.
XX
PS    Claim 2; SEQ ID NO 86; 591pp; English.
XX
CC    The invention relates to novel human diagnostic and therapeutic
CC    polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded
CC    proteins (DITHP; ABR41136-ABR41812). The invention also relates to
CC    polynucleotide sequences at least 90% identical to the dithp cDNA
CC    sequences of the invention; recombinant vectors, host cells and
CC    transgenic organisms comprising a dithp nucleic acid sequence; the
CC    recombinant production of DITHP proteins; antibodies specific for DITHP
CC    proteins; microarrays comprising dithp nucleic acid sequences; methods of
CC    detecting dithp nucleotide and protein sequences; methods of screening
CC    for compounds which specifically bind a DITHP protein; and methods of
CC    assessing the toxicity of test compounds using a dithp hybridisation
CC    probe. Dithp nucleic acid sequences and DITHP proteins may be used in the
CC    diagnosis of a wide variety of conditions including cancer and other cell
CC    proliferative disorders; autoimmune or inflammatory disorders; bacterial,
CC    viral, fungal or parasitic infections; hormonal disorders; metabolic
CC    disorders; neurological disorders; gastrointestinal disorders; transport
CC    disorders; and connective tissue disorders. They may also be used to
CC    screen for modulators of protein activity or gene expression. DITHP
CC    proteins can additionally be used in analysis of the proteome of a tissue
CC    or cell type and to induce antibodies. The dithp nucleic acids are
```

SQ   Sequence 1473 BP; 403 A; 320 C; 303 G; 447 T; 0 U; 0 Other;

       Query Match   38.4%;   Score 592.4;   DB 7;   Length 1473;
       Best Local Similarity 75.3%;   Pred. No. 1.6e-139;
       Matches 764; Conservative   0; Mismatches 246;   Indels   4; Gaps   2;

Qy   39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
            | | ||||||| |||| | |||| | || ||| ||||| ||||||||| | || | |
Db   119 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 178

Qy   99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
            |||||||||| || ||||||| || ||||||| || | ||| | ||||| ||| ||
Db   179 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 238

Qy   159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
            ||| || | ||||||| |||||| ||||| |||||||||||||||| ||| | |||||| |
Db   239 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 298

Qy   219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
            ||||||| || |||||||||| ||||| |||||||||||| ||||| ||||||| |||||
Db   299 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 358

Qy   279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
            ||||||||| || ||| ||||||||| || ||||| |||||||||||||||||||||||||||||
Db   359 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 418

Qy   339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
            | ||||||||||| ||||||||| |||| | |||||||| || ||||| || ||||| ||
Db   419 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 478

Qy   399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
            | || |||||| |||||||||||||||| |||| |||||||| || ||||| ||||||||||
Db   479 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCT 538

Qy   459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
            | |||| | |||| ||||| ||||||||| | |||||||| || | | | || |||
Db   539 CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 598

Qy   519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
            |||| || || | |||| | ||| | || | ||||||||||||||||| |||| |
Db   599 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 658

Qy   579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
            |||| |||||||||||| | || || || ||||||| ||||| |||||||| | |||||
Db   659 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 718

```
Qy    639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
          ||||  ||||||  ||  |||||||||  |    |||||  |||||   |||||     ||||   || |
Db    719 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 778

Qy    699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
          ||||  ||||||  ||  ||  ||  || |     |  ||||  || ||||  ||  ||  |||||||||
Db    779 CTGCTCTGCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 838

Qy    759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
          |  ||  ||  ||||||||||||  |||||||| |||  |||||||||||| ||||||||||  ||||
Db    839 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 898

Qy    819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
          |   ||    |    || ||  |||    |  |||||| ||  |  ||||       |||||||||||||||
Db    899 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT 958

Qy    876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
          |||||||||||||||||||||  |||||||  ||   ||||||  || ||  |||||||  ||  | ||
Db    959 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 1018

Qy    936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
          | ||  ||||||||| |  |  |  ||||||  ||||||||  ||||||||||||||| ||  |  ||
Db   1019 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 1078

Qy    996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
          | |||  || |  |      ||||  ||||  ||||  ||||  ||||||  ||||||
Db   1079 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAAGTGAGGGGCTTGTGAAACAG 1132
```

RESULT 8
AAD24958
ID   AAD24958 standard; cDNA; 1542 BP.
XX
AC   AAD24958;
XX
DT   12-MAR-2002   (first entry)
XX
DE   Human G-protein coupled receptor-3 (GCREC-3) cDNA.
XX
KW   Human; G-protein coupled receptor-3; GCREC-3; therapy; cancer; stroke;
KW   cell proliferative disorder; neurological; epilepsy; Parkinson's disease;
KW   Alzheimer's disease; inflammation; thyroiditis; haemolytic anaemia; AIDS;
KW   Acquired Immune Deficiency Syndrome; dementia; nootropic; cholelithiasis;
KW   multiple sclerosis; atherosclerosis; angina pectoris; gastroenteritis;
KW   diabetes; ulcer; viral infection; immunosuppressive; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   63. .1202
FT   /*tag= a
FT   /product= "Human GCREC-3 protein"
XX
PN   WO200198351-A2.
XX
PD   27-DEC-2001.
XX

PT   Isolated human G-protein coupled receptor polypeptides and the use of
PT   these sequences in the diagnosis, treatment and prevention of diseases
PT   and in the assessment of exogenous compounds on the expression of the
PT   receptors.
XX
PS   Claim 11; Page 133; 143pp; English.
XX
CC   The invention relates to isolated human G-protein coupled receptor
CC   (GCREC) polypeptides and their biologically active fragments. GCREC and
CC   protein is useful in treating a disease or condition associated with an
CC   increase or decrease in expression of functional GCREC. The GCREC's are
CC   useful in the diagnosis, treatment and prevention of cell proliferative
CC   disorders (cancer, leukaemia, melanoma); neurological disorders (stroke,
CC   epilepsy, Parkinson's disease, dementia, Alzheimer's disease); autoimmune
CC   inflammatory disorder (thyroiditis, haemolytic anaemia, AIDS, multiple
CC   sclerosis); cardiovascular disorder (atherosclerosis, angina pectoris),
CC   gastrointestinal disorder (ulcer, cholelithiasis, gastroenteritis),
CC   metabolic disorders (diabetes); viral infections (herpes virus) and in
CC   the assessment of the effects of exogenous compounds on the expression of
CC   the nucleic acid and amino acid sequences. The present sequence is human
CC   GCREC-3 cDNA
XX
SQ   Sequence 1542 BP; 428 A; 327 C; 315 G; 472 T; 0 U; 0 Other;

      Query Match   38.4%; Score 592.4; DB 6; Length 1542;
      Best Local Similarity   75.3%; Pred. No. 1.6e-139;
      Matches 764; Conservative   0; Mismatches 246; Indels   4; Gaps   2;

Qy   39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
            | | ||||||| |||| | |||| | || ||| |||||| |||||||| | || | |
Db   205 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 264

Qy   99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
            |||||||||| || ||||||| || |||||| || | ||| | ||||| ||| ||
Db   265 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 324

Qy   159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218

```
           |||  ||  |  |||||||  ||||||    ||||||||||||||||||||  |||  |  ||||||  |
Db    325 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 384

Qy    219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
          |||||||  ||  |||||||||||  |||||  |||||||||  |||||  |||||||  |||||
Db    385 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 444

Qy    279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
          ||||||||  ||   |||  ||||||||||  ||  |||||  ||||||||||||||||||||||||
Db    445 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 504

Qy    339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
          |  ||||||||||  |||||||  |||||  |||||  ||||||  ||  |||||  ||  |||||  ||
Db    505 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 564

Qy    399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAGAAGGAATTTGCCATTTTAATCT 458
          |  ||  ||||||  |||||||||||||||||||  ||||  ||||||||  ||||||  ||||||||||
Db    565 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAGAAAGAGTTTGCTATTTTAATCT 624

Qy    459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
          |   ||||  |  ||||  |||||  ||||||||  |  |||||||||  ||  |  |  ||  |||
Db    625 CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 684

Qy    519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
          ||||     ||   ||   |  ||||   |  |||  |   ||  |  |||||||||||||  ||||   |
Db    685 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTGCAAGTTCTGGAGACCCCAACT 744

Qy    579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
          ||||  ||||||||||||  |  ||  ||  ||  |||||||  |||||  ||||||||  |  |||||
Db    745 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 804

Qy    639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
          ||||  |||||  ||  |||||||||  |      |||||  |||||   |||||   ||  |
Db    805 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 864

Qy    699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
          ||||  |||||  ||  ||  ||  ||  ||  |      |  ||||  ||  ||||  ||  ||  |||||||||
Db    865 CTGCTCTGCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 924

Qy    759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
          |  ||  ||  ||||||||||||  |||||||  |||||||  |||  ||||||||||||||||  ||||
Db    925 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 984

Qy    819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
          |   ||   |     ||  ||  |||   |  ||||||  ||  |  ||||       |||||||||||
Db    985 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT 1044

Qy    876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
          |||||||||||||||||||||||  |||||||  ||  |||||||  ||  ||  ||||||||  ||  |  |||
Db   1045 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 1104

Qy    936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
          |  ||  ||||||||  |  |  |  |  ||||||  |||||||  ||||||||||||||||||  ||  |  |||
Db   1105 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 1164

Qy    996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
          |  |||  ||  |  |   ||||  ||||  ||||  ||||    ||||||  ||||||
```

RESULT 9
ABS57291
ID   ABS57291 standard; cDNA; 1338 BP.
XX
AC   ABS57291;
XX
DT   30-JAN-2003   (first entry)
XX
DE   cDNA encoding human adenosine receptor.
XX
KW   Human; mammalian; adenosine receptor; G-protein coupled receptor; GPCR;
KW   adenosine-mediated medical condition; vasodilation; hypotension;
KW   reversal of tachycardia; chronic renal disease; thyroid disorder;
KW   inflammation; asthma; hypertensive; antiarrhythmic; antiinflammatory;
KW   antiasthmatic; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1. .1005
FT   /*tag= a
FT   /product= "Adenosine receptor"
XX
PN   US2002137887-A1.
XX
PD   26-SEP-2002.
XX
PF   17-JAN-2001; 2001US-00765034.
XX
PR   17-JAN-2001; 2001US-00765034.
XX
PA   (HEDR/) HEDRICK J A.
PA   (LACH/) LACHOWICZ J E.
PA   (WANG/) WANG W.
PA   (GUST/) GUSTAFSON E L.
XX
PI   Hedrick JA, Lachowicz JE, Wang W, Gustafson EL;
XX
DR   WPI; 2003-074992/07.
DR   P-PSDB; ABG72131.
XX
PT   Novel isolated mammalian adenosine receptor polypeptide useful for
PT   identifying an agonist or antagonist of the receptor for treating
PT   vasodilation, hypotension, chronic renal diseases, thyroid disorders and
PT   inflammation.
XX
PS   Example 1; Page 14-16; 19pp; English.
XX
CC   The present invention relates to the isolation of a mammalian (human)
CC   adenosine receptor, and the polynucleotide sequence encoding it. The
CC   cloned receptor resembles a member of the G-protein coupled receptor
CC   (GPCR) superfamily that contains 7-transmembrane domains. The adenosine
CC   receptor is useful for identifying agonists and antagonists of the
CC   receptor, which may be useful for treating an adenosine-mediated medical

CC   condition. The adenosine receptor polypeptide sequence is also useful as
CC   an antigen to elicit antibody production in an immunologically competent
CC   host. An antibody which binds specifically to the adenosine receptor is
CC   useful for treating medical conditions caused or mediated by adenosine
CC   such as vasodilation, hypotension, reversal of tachycardia, chronic renal
CC   diseases, thyroid disorders and inflammation (e.g. asthma). The antibody
CC   can also be used to purify the adenosine receptor, or as a basis for
CC   immunoassays of the receptor. The polynucleotide sequence encoding the
CC   adenosine receptor is useful for producing vectors and host cells
CC   containing the vectors. It is also useful for measuring expression of a
CC   mammalian adenosine receptor gene in a biological sample. The present
CC   sequence encodes human adenosine receptor
XX
SQ   Sequence 1338 BP; 370 A; 288 C; 265 G; 415 T; 0 U; 0 Other;

      Query Match   38.3%;   Score 590.8;   DB 7;   Length 1338;
      Best Local Similarity 75.2%;   Pred. No. 3.9e-139;
      Matches 763; Conservative   0; Mismatches 247; Indels   4; Gaps   2;

Qy   39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
            | | |||||| |||| | |||| | || ||| ||||| |||||||| | || | | |
Db   8 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 67

Qy   99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
            ||||||||||| || ||||||| || |||||| || | ||| | ||||| || ||
Db   68 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 127

Qy   159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
            ||| || | ||||||| |||||| ||||||||||||||||||| ||| |||||| |
Db   128 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 187

Qy   219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
            ||||||| || |||||||||| ||||| ||||||||||| ||||| ||||||| | |||||
Db   188 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 247

Qy   279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
            ||||||||| || ||| ||||||||| || ||||| ||||||||||||||||||||||||
Db   248 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 307

Qy   339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
            | |||||||||| |||||||| ||||| |||||||| || ||||| || ||||| ||
Db   308 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 367

Qy   399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
            | || |||||| |||||||||||||||||||| |||| ||||||||| ||||| ||||||||||
Db   368 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCT 427

Qy   459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
            | |||| | |||| ||||| |||||||||| | ||||||||| | | | || |||
Db   428 CCTTGGCCATTTGGGTTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 487

Qy   519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
            |||| || || | | |||| | ||| | || | |||||||||||||||| |||| |
Db   488 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 547

Qy   579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
            |||| ||||||||||||| | || || || || ||||||||| ||||| |||||||| | |||||

```
Db    548 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 607

Qy    639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
          ||||  ||||||  ||  ||||||||||  |   |||||| ||||| |||||    ||||    || |
Db    608 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 667

Qy    699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
          ||||  |||||| ||  || || || |      |  ||||  ||  ||||  ||  ||  |||||||||||
Db    668 CTGCTCTGCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 727

Qy    759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
          | |  ||  |||||||||||   |||||||  ||| ||||||||||| |||||||||||   ||||
Db    728 TGCCTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 787

Qy    819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACTGACACGGCCTC 875
          |   ||   |    || || |||    | |||||| ||  | ||||    |||||||||||||
Db    788 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT 847

Qy    876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
          ||||||||||||||||||||| |||||| ||  |||||||  || ||  ||||||| ||  | ||
Db    848 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 907

Qy    936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
          | ||  |||||||||| | | |  | |||||| ||||||| |||||||||||||||||| || | ||
Db    908 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 967

Qy    996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
          | |||  || |  |     ||||  |||| |||| ||||  ||||  ||||||
Db    968 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAGTGAGGGGCTTGTGAAACAG 1021
```

RESULT 10
ACD27619
ID   ACD27619 standard; cDNA; 1428 BP.
XX
AC   ACD27619;
XX
DT   18-SEP-2003   (first entry)
XX
DE   Human ATP receptor cDNA.
XX
KW   Human; ss; gene; ATP receptor; G-protein coupled receptor; gene therapy;
KW   7-transmembrane receptor; asthma; allergic rhinitis; hypertension; ulcer;
KW   angina pectoris; allergy; psychosis; depression; migraine; vomiting;
KW   benign prostatic hypertrophy; arterial thrombosis; myocardial infarction;
KW   urinary retention; angioplasty; cystic fibrosis; Parkinson's disease;
KW   acute heart failure; hypotension; thrombolysis; osteoporosis.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   91. .1096
FT   /*tag= a
FT   /product= "ATP receptor"
XX
PN   US2003054487-A1.
XX

PT   New polynucleotide, useful for producing a medicament for treating
PT   asthma, allergic rhinitis or hypertension.
XX
PS   Claim 1; Fig 1; 24pp; English.
XX
CC   The invention relates to an isolated polynucleotide encoding a G-protein
CC   coupled, 7-transmembrane ATP receptor. The polynucleotide is useful for
CC   producing a medicament for treating asthma, allergic rhinitis or
CC   hypertension. Antagonists for the the ATP receptor can be used to treat
CC   angina pectoris, ulcers, allergies, psychoses, depression, migraine,
CC   vomiting, benign prostatic hypertrophy, arterial thrombosis, myocardial
CC   infarction, thrombolysis, angioplasty, cystic fibrosis. Agonists of the
CC   ATP receptor can be used to treat Parkinson's disease, acute heart
CC   failure, hypotension, urinary retention and osteoporosis. The present
CC   sequence represents cDNA encoding the human ATP receptor
XX
SQ   Sequence 1428 BP; 394 A; 306 C; 290 G; 438 T; 0 U; 0 Other;

    Query Match   38.3%;   Score 590.8;   DB 8;   Length 1428;
    Best Local Similarity   75.2%;   Pred. No. 4e-139;
    Matches   763;   Conservative   0;   Mismatches   247;   Indels   4;   Gaps   2;

Qy   39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
            | | ||||||| |||| | |||| | || ||| ||||| ||||||||| | ||| | |
Db   99 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 158

Qy   99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
            ||||||||||| || ||||||| || |||||| || | ||| | |||||| ||| ||
Db   159 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 218

Qy   159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
            ||| || | ||||||| |||||| |||||||||||||||||||| ||| | |||||| |
Db   219 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 278

Qy   219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
            ||||||| || ||||||||||| ||||| |||||||||||| ||||| ||||||||| |||||
Db   279 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 338

Qy   279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
            ||||||||| || ||| ||||||||||| || ||||| ||||||||||||||||||||||||||||||||||||
Db   339 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 398

```
Qy     339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
           |  |||||||||| |||||||| ||||| |||||||| || ||||| || ||||| ||
Db     399 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 458

Qy     399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
           | || |||||| |||||||||||||||||| |||| ||||||||| || | ||| ||||||||||
Db     459 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAGAAAGAGTGTGCTATTTTAATCT 518

Qy.    459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
           | |||| | |||| ||||| ||||||||| | ||||||||| || | | || |||
Db     519 CCTTGGCCATGTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 578

Qy     519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
           ||||    ||  ||  | |||| | |||| |  || | |||||||||||||||| ||||  |
Db     579 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 638

Qy     579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
           |||| |||||||||||| | || || || || ||||||| ||||| |||||||| | |||||
Db     639 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 698

Qy     639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
           |||| ||||| || ||||||||| |     ||||| ||||| ||||| |||||     |||| || |
Db     699 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 758

Qy     699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
           |||| |||||| || || || || || |    | |||| || |||| || || ||||||||||
Db     759 CTGCTCTGCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 818

Qy     759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
           | || || ||||||||||| |||||||| | |||||||||||| ||||||||||||| ||||
Db     819 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 878

Qy     819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
           |   || |      || || ||| | |||||| || | |||| |||||||||||
Db     879 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTG 938

Qy     876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
           |||||||||||||||||| |||||| || ||||||| || || ||||||| || | ||
Db     939 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTGTGGGAGATCACTTCA 998

Qy     936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
           | || ||||||||| | | |  | ||||| ||||||| |||||||||||||||||| || | ||
Db     999 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 1058

Qy     996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
           | ||| || |   |     |||| |||| |||| ||||    |||| ||||||
Db    1059 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAAGTGAGGGGCTTGTGAAACAG 1112
```

RESULT 11
AAT71900
ID   AAT71900 standard; cDNA; 1996 BP.
XX
AC   AAT71900;
XX
DT   11-SEP-1997 (first entry)
XX

```
DE    Human purinergic receptor P2U2 cDNA.
XX
KW    P2U2 receptor; purinergic receptor; diagnosis; therapy; ss.
XX
OS    Homo sapiens.
XX
FH    Key              Location/Qualifiers
FT    CDS              625. .1629
FT                     /*tag=  a
XX
PN    WO9720045-A2.
XX
PD    05-JUN-1997.
XX
PF    08-NOV-1996;     96WO-US018175.
XX
PR    15-NOV-1995;     95US-0006782P.
PR    15-NOV-1995;     95US-00559524.
XX
PA    (CORT-) COR THERAPEUTICS INC.
XX
PI    Conley PB,  Jantzen H;
XX
DR    WPI; 1997-310601/28.
DR    P-PSDB; AAW19854.
XX
PT    New isolated purinergic receptor sub-type - used to develop products for
PT    diagnosis and therapy, e.g. for screening for agonists and antagonists
PT    which can modulate activation.
XX
PS    Claim 3; Fig 1A-C; 36pp; English.
XX
CC    A cDNA clone (AAT71900) codes for a novel human purinergic receptor
CC    subtype, designated P2U2 receptor (AAW19854), that is abundantly
CC    expressed in kidney and in many cell lines of megakaryocytic or
CC    erythroleukaemic origin and which is activated by ATP, UDP, UTP and UDP.
CC    The clone was obtd. by amplifying DAMI (ATCC CRL 9792) cell cDNA using
CC    primers (see also AAT72104-05) based on transmembrane regions of mouse
CC    P2u and chicken P2Y1 receptors, and use of the PCR product to screen the
CC    DAMI cDNA library to isolate the full-length clone. P2U2 nucleic acids
CC    can be used in the recombinant prodn. of P2U2 receptor polypeptides and
CC    as probes
XX
SQ    Sequence 1996 BP; 513 A; 454 C; 381 G; 647 T; 0 U; 1 Other;

   Query Match              38.2%;  Score 589.2;  DB 2;  Length 1996;
   Best Local Similarity    75.1%;  Pred. No. 1.2e-138;
   Matches  762;  Conservative   0;  Mismatches  248;  Indels    4;  Gaps    2;

Qy        39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
             | |   |||||||  ||||  | ||||  | || ||| |||||| ||||||||| | || | |
Db       632 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 691

Qy        99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
             |||||||||| ||     |||||||  || ||||||| || | ||| | ||||| |||  ||
Db       692 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 751
```

```
Qy   159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
         ||| || | ||||||| ||||||    |||||||||||||||||||||||| ||| | |||||| |
Db   752 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 811

Qy   219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
         |||||||| || ||||||||||| ||||| ||||||||||| ||||| ||||||| |||||
Db   812 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 871

Qy   279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
         |||||||||| || ||| ||||||||| || ||||| |||||||||||||||||||||||||
Db   872 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 931

Qy   339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
         |  ||||||||| |||||||| ||||||| ||||||| || ||||| || ||||| ||
Db   932 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 991

Qy   399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
         | || ||||| |||||||||||||||| |||| |||||||| || ||||| |||||||||||
Db   992 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAGAAAGAGTTTGCTATTTTAATCT 1051

Qy   459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
         |  |||| | |||| ||||| |||||||| | ||||||||| || |  | ||  |||
Db  1052 CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 1111

Qy   519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
         ||||    || || |  |||| | ||| | || | |||||||||||||||| ||||  |
Db  1112 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 1171

Qy   579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTGTGA 638
         |||| |||||||||||| ||| || | || || || ||||||| ||||| |||||||| | |||||
Db  1172 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTGTGA 1231

Qy   639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
         |||| ||||| || ||||||||| |     ||||| |||||     ||||     |||| || |
Db  1232 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 1291

Qy   699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
         |||| |||||| || || || || |    | |||| || ||||| || || |||||||||
Db  1292 CTGCTCTGCCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 1351

Qy   759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
         | || || |||||||||| || ||||| |||| |||||||| |||| || |||||| ||||||
Db  1352 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 1411

Qy   819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
         |   ||   |     || || ||| | |||||| || | |||| ||||||||| ||
Db  1412 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGGCTT 1471

Qy   876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
         ||| ||||||||||||||||| ||||||| |||||| || |||||||| || || |||||| || | ||
Db  1472 TGGGCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 1531

Qy   936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
         | || |||||||||| | | | | ||||| |||||| ||||||||||||||||||| || | ||
Db  1532 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 1591

Qy   996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
```

```
RESULT 12
AAT75146
ID    AAT75146 standard; cDNA; 1428 BP.
XX
AC    AAT75146;
XX
DT    07-OCT-1997  (first entry)
XX
DE    Human ATP receptor cDNA.
XX
KW    ATP receptor; G-protein coupled receptor; agonist; antagonist; ss.
XX
OS    Homo sapiens.
XX
FH    Key             Location/Qualifiers
FT    CDS             92. .1096
FT                    /*tag=  a
FT                    /transl_except= (pos:725. .727, aa:Ser)
FT                    /transl_except= (pos:764. .766, aa:Ser)
FT                    /transl_except= (pos:820. .822, Xaa)
FT                    /note= "Xaa = unknown"
FT    primer_bind     complement(92. .109)
FT                    /*tag=  c
FT                    /note= "binding site for primers used to amplify cDNA for
FT                    bacterial or COS expression"
FT    primer_bind     complement(92. .100)
FT                    /*tag=  b
FT                    /note= "binding site for primer used to amplify cDNA for
FT                    baculovirus expression"
FT    primer_bind     1076. .1095
FT                    /*tag=  d
FT                    /note= "binding site for primer used to amplify cDNA for
FT                    COS expression"
FT    primer_bind     1079. .1096
FT                    /*tag=  e
FT                    /note= "binding site for primer used to amplify cDNA for
FT                    bacterial expression"
FT    primer_bind     1085. .1096
FT                    /*tag=  f
FT                    /note= "binding site for primer used to amplify cDNA for
FT                    baculovirus expression"
XX
PN    WO9724929-A1.
XX
PD    17-JUL-1997.
XX
PF    11-JAN-1996;    96WO-US000392.
XX
PR    11-JAN-1996;    96WO-US000392.
XX
PA    (HUMA-) HUMAN GENOME SCI INC.
XX
PI    Li Y;
```

PT   Isolated human ATP receptor - agonists and antagonists of which are
PT   useful in treatment of, e.g. asthma, hypertension, arterial thrombosis
PT   and psychotic and neurological disorders.
XX
PS   Claim 7; Fig 1A-C; 53pp; English.
XX
CC   A cDNA clone (AAT75146) codes for human ATP receptor (AAW22732), a
CC   polypeptide structurally related to the G protein-coupled receptor
CC   family. It was discovered in a human placenta cDNA library. cDNA encoding
CC   the mature receptor, deposited as ATCC 97333, can be expressed in
CC   bacterial (e.g. E. coli), mammalian (e.g. COS) or insect (e.g. Sf9) host
CC   cells and used to screen for agonists and antagonists useful in the
CC   treatment of a variety of disorders. It can also be used to identify a
CC   mutation in an ATP receptor gene and thus to diagnose diseases, or
CC   susceptibility to diseases, related to ATP receptor underexpression
XX
SQ   Sequence 1428 BP; 394 A; 308 C; 290 G; 435 T; 0 U; 1 Other;

        Query Match   38.1%;   Score 587.2;   DB 2;   Length 1428;
        Best Local Similarity   75.0%;   Pred. No. 3.2e-138;
          Matches 760; Conservative   1; Mismatches 249; Indels   4; Gaps   2;

```
Qy         39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
              | |   ||||||  ||||  |  ||||  | || |||  |||||  |||||||| | ||  | |
Db         99 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 158

Qy         99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
              ||||||||||| ||     |||||||  || |||||| || | ||| | ||||| |||  ||
Db        159 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 218

Qy        159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
              ||| || | |||||||| ||||||       ||||||||||||||||||| ||| | |||||| |
Db        219 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 278

Qy        219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
              |||||||  || |||||||||| |||||  |||||||||||  |||||  |||||||| |||||
Db        279 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 338

Qy        279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
              |||||||||  ||  |||   ||||||||| || ||||| |||||||||||||||||||||||||||
Db        339 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 398

Qy        339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
              |   |||||||||| ||||||||| |||||  ||||||||| || ||||| || ||||||  ||
Db        399 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 458

Qy        399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
              | || ||||| ||||||||||||||||| |||| |||||||| || | ||| ||||||||||
Db        459 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAGAAAGAGTGTGCTATTTTAATCT 518

Qy        459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
              |  ||||  |  ||||| |||||  ||||||||| | |||||||| || |  |  || |||
Db        519 CCTTGGCCATGTGGGTTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 578
```

```
Qy      519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
                ||||   ||  ||  | |||| | ||| |  || | |||||||||||||| |||| |
Db      579 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 638

Qy      579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTGTGA 638
                ||||  ||||||||||||| | ||  ||  ||   ||||||| |||||  |||||||| | |||||
Db      639 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 698

Qy      639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
                |||| ||||| || |||||||| |    |||| ||||| ||||| |||||    |||| || |
Db      699 TGTGTTTCTTTTATTACAAGATTGCCTCCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 758

Qy      699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
                ||||| ||| || || || || || |    | |||| || |||| || || |||||||||
Db      759 CTGCCTCGCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 818

Qy      759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
              | |: || ||||||||||||  |||||||| ||| |||||||||| |||||||||  ||||
Db      819 TGCYTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 878

Qy      819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
              |   ||  |    ||  || ||| |  |||||| || | |||    |||||||||||
Db      879 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTG 938

Qy      876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
              |||||||||||||||||||| |||||| |||||||| || |||||||| || | ||
Db      939 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTGTGGGAGATCACTTCA 998

Qy      936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
              | || |||||||| | | | | ||||| ||||||| |||||||||||||| || | ||
Db      999 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 1058

Qy      996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
              | ||| || | |    |||| |||| |||| ||||  |||| ||||||
Db     1059 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAGTGAGGGGCTTGTGAAACAG 1112
```

RESULT 13
AAC81122
ID   AAC81122 standard; cDNA; 1385 BP.
XX
AC   AAC81122;
XX
DT   14-FEB-2001 (first entry)
XX
DE   Human secreted protein gene 37 SEQ ID NO:47.
XX
KW   Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
KW   antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
KW   cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
KW   fungicide; ophthalmological; vulnerary; gene therapy; autoimmune disease;
KW   hyperproliferative disorder; cardiovascular disorder; angiogenesis;
KW   cerebrovascular disorder; nervous system disorder; infection; skin aging;
KW   ocular disorder; wound healing; food additive; preservative; ss.
XX
OS   Homo sapiens.

CC   The polynucleotide sequences given in AAC81086 to AAC81134 encode the
CC   human secreted proteins given in AAB45308 to AAB45356. AAB45357 to
CC   AAB45384 represent human secreted polypeptide sequences and proteins
CC   homologous to them, which are given in the exemplification of the present
CC   invention. Human secreted proteins have activities based on the tissues
CC   and cells the genes are expressed in. Examples of activities include:
CC   antiarthritic; immunosuppressive; antirheumatic; antiproliferative;
CC   cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;
CC   neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
CC   and vulnerary. The polynucleotides and polypeptides can be used to
CC   prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC   rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC   in diagnosing a pathological condition or susceptibility to a
CC   pathological condition. Disorders which are diagnosed or treated include
CC   autoimmune diseases, hyperproliferative disorders, cardiovascular
CC   disorders, cerebrovascular disorders, angiogenesis, nervous system
CC   disorders, infections caused by bacteria, viruses and fungi and ocular
CC   disorders. The polypeptides can also be used to aid wound healing and
CC   epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC   maintain organs before transplantation, for supporting cell culture of
CC   primary tissues, to regenerate tissues and in chemotaxis. The
CC   polypeptides can also be used as a food additive or preservative to
CC   increase or decrease storage capabilities, fat content, lipid, protein,
CC   carbohydrate, vitamins, minerals, cofactors and other nutritional
CC   components. AAC81077 to AAC81085 and AAB45307 represent sequences used in
CC   the exemplification of the present invention
XX
SQ   Sequence 1385 BP; 385 A; 296 C; 275 G; 429 T; 0 U; 0 Other;

     Query Match   37.6%;   Score 580.4;   DB 3;   Length 1385;
     Best Local Similarity   75.2%;   Pred. No. 1.7e-136;
     Matches   763; Conservative   0; Mismatches   246;   Indels   5;   Gaps   3;

Qy   39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98

```
            | |   ||||||   ||||   | ||||   | ||  |||  |||||  ||||||||  |  || | |
Db    56 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 115

Qy    99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
         |||||||||| ||    ||||||||   ||  ||||||| ||  | |||  |  ||||| |||   ||
Db   116 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 175

Qy   159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
         |||  ||   |||||||  ||||||    ||||||||||||||||||||||| ||| | ||||||  |
Db   176 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 235

Qy   219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
         ||||||| ||  |||||||||| ||||| ||||| |||||||||| ||||| ||||||| |||||
Db   236 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 295

Qy   279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
         |||||||| || |||   ||||||| || |||||  ||||||||||||||||||||||||||||
Db   296 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 355

Qy   339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
         |  |||||||||| ||||||||| ||||||| || ||||||| || |||||| || |||||| ||
Db   356 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 415

Qy   399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
         | ||  |||||  |||||||||||||||||||| |||  |||||||| || ||||| ||||||||||
Db   416 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCT 475

Qy   459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
         |  ||||  | ||||  |||||  ||||||||  |  ||||||||| ||  |  | || |||
Db   476 CCTTGGCCATTTGGGTTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 535

Qy   519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
         ||||     ||   ||   | ||||   | |||  || | |||||||||||||||| ||||  |
Db   536 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 595

Qy   579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
         ||||  ||||||||||||  ||  | || || || |||||||| ||||| |||||||| | |||||
Db   596 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 655

Qy   639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
         ||||  |||||  || |||||||  | ||||| ||||| |||||  |||||    ||||  || |
Db   656 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 715

Qy   699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
         ||||  ||||| ||  || || ||  |  | |||| ||  ||||  || ||  ||||||||||
Db   716 CTGCTCTGCCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 775

Qy   759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
         | || || |||||||||||| | ||||| ||||||| |||||||||| ||||||||||||| ||||
Db   776 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 835

Qy   819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
         |   ||  | | ||  || ||| ||  | |||||||| ||  |  ||||    |||||||||||
Db   836 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACAC-GCCTT 894

Qy   876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
         ||||||||||||||||||||||| ||||||| ||  ||||||| ||  || |||||||| || | ||
```

```
Db        895 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 954

Qy        936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
              | || |||||||| | | | | ||||| ||||||| ||||||||||||||| || | ||
Db        955 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 1014

Qy        996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
              | ||| || | |    |||| |||| |||| |||| |||| ||||||
Db       1015 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAAGTGAGGGGCTTGTGAAACAG 1068
```

RESULT 14
ADC12679

ID   ADC12679 standard; DNA; 1005 BP.
XX
AC   ADC12679;
XX
DT   18-DEC-2003   (first entry)
XX
DE   Human GPCR gene, SEQ ID No 11.
XX
KW   G protein-coupled receptor; GPCR; antibacterial; fungicide; protozoacide;
KW   virucide; antirheumatic; antiarthritic; tranquiliser; antidiabetic;
KW   osteopathic; nootropic; neuroprotective; anorectic; cardiant;
KW   neuroleptic; cytostatic; antiparkinsonian; hypotensive; hypertensive;
KW   antiulcer; antiallergic; anticonvulsant; analgesic; infection;
KW   rheumatoid arthritis; chronic obstructive pulmonary diseases; COPD;
KW   asthma; non-insulin dependent diabetes; obesity; osteoporosis;
KW   Alzheimer's disease; age-related macular degeneration;
KW   myocardial infarction; schizophrenia; osteoarthritis; cancer;
KW   Parkinson's disease; congestive heart failure; hypotension; hypertension;
KW   ulcer; allergy; benign prostatic hyperplasia; seizure disorder; anxiety;
KW   obsessive compulsive disorder; Cushing's syndrome; hypopituitarism; pain;
KW   gene; ds; human.
XX
OS   Homo sapiens.
XX
PN   WO2003000893-A2.
XX
PD   03-JAN-2003.
XX
PF   24-JUN-2002; 2002WO-IB002357.
XX
PR   26-JUN-2001; 2001US-0301095P.
PR   06-NOV-2001; 2001US-0333185P.
XX
PA   (DECO-) DECODE GENETICS EHF.
XX
PI   Martinez RMA, Sigurdsson GT;
XX
DR   WPI; 2003-210155/20.
DR   P-PSDB; ADC12680.
XX
PT   New G protein-coupled receptor (GPCR) genes and polypeptides, useful for
PT   diagnosing diseases associated with a GPCR, or in gene therapy for
PT   treating e.g. obesity, osteoporosis, Alzheimer's, cancers or congestive
PT   heart failure.
```

CC    The invention relates to a novel isolated nucleic acid of a G protein-
CC    coupled receptor (GPCR) gene comprising any of 62 sequences of 912-2454
CC    bp, or its complements; a GPCR polypeptide comprising any of 62 sequences
CC    of 291-818 amino acids; or a nucleic acid that hybridises, under high
CC    stringency conditions, with any of the 62 GPCR sequences or any of their
CC    complements. The GPCR agents of the invention have the following
CC    activities: antibacterial, fungicide, protozoacide, virucide,
CC    antirheumatic, tranquiliser, antiarthritic, antidiabetic, osteopathic,
CC    nootropic, neuroprotective, anorectic, cardiant, neuroleptic, cytostatic,
CC    antiparkinsonian, hypotensive, hypertensive, antiulcer, antiallergic,
CC    anticonvulsant, and analgesic. The GPCR therapeutic agent, particularly a
CC    GPCR gene agonist or antagonist, is useful for treating a disease or
CC    condition associated with a GPCR in an individual. The nucleic acid cited
CC    above, which is 100 or fewer nucleotides in length, is useful for
CC    assaying a sample for the presence of the GPCR gene nucleic acid or a
CC    GPCR gene nucleic acid with at least one nucleotide difference from a
CC    first nucleic acid, or for diagnosing a susceptibility to a disease or
CC    conditions associated with a GPCR. These diseases include infections
CC    (e.g. bacterial, fungal, protozoan or viral), rheumatoid arthritis,
CC    chronic obstructive pulmonary diseases (COPD), asthma, non-insulin
CC    dependent diabetes, obesity, osteoporosis, Alzheimer's disease, age-
CC    related macular degeneration, myocardial infarction, schizophrenia,
CC    osteoarthritis, cancers, Parkinson's diseases, congestive heart failure,
CC    hypotension, hypertension, ulcers, allergies, benign prostatic
CC    hyperplasia, seizure disorder, anxiety, obsessive compulsive disorder,
CC    Cushing's syndrome, hypopituitarism, or pain. This polynucleotide
CC    sequence represents one of the 62 GPCR gene sequences of the invention.
XX
SQ    Sequence 1005 BP; 244 A; 246 C; 187 G; 328 T; 0 U; 0 Other;

      Query Match               37.1%;  Score 572.8;  DB 9;  Length 1005;
      Best Local Similarity     74.6%;  Pred. No. 1.2e-134;
      Matches   734;  Conservative     0;  Mismatches   247;  Indels    3;  Gaps    1;

Qy        86 GCTATCTTGAATAAGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTG 145
             ||| | || | |||||||||||| ||   ||||||| || |||||| || | ||| |
Db         1 GCTGCCCTGGAAAAGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTC 60

Qy       146 CTTGGGAATGTCACTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGC 205
             ||||| ||| || ||| || | ||||||| |||||| ||||||||||||||||||||||
Db        61 CTTGGAAATACCATTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGT 120

Qy       206 AATGTCTATCTTTTTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATC 265
             ||| | |||||| ||||||||| || |||||||||| ||||| ||||||||||| |||||
Db       121 AATATTTATCTCTTTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATG 180

Qy       266 CTGATAAAGAGTTATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAAC 325
             ||||||| |||||||||||||||| || ||| ||||||||| || ||||| |||||||||
Db       181 CTGATAAGGAGTTATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAAC 240

Qy       326 CGATATGTGCTTCACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATG 385
             ||||||||||||| ||||||||||| ||||||||| || ||||| ||||||| || |||||
Db       241 CGATATGTGCTTCATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATA 300

```
Qy    386 GACCGATATCTGCTCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTT 445
          || ||||| || | || ||||| |||||||||||||||| |||| |||||||||| || |||
Db    301 GATCGATACTTGATAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTT 360

Qy    446 GCCATTTTAATCTCGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTC 505
          || ||||||||||| |||| | |||| ||||| ||||||||| | |||||||| ||
Db    361 GCTATTTTAATCTCCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTT 420

Qy    506 ACTTTCATCAATTCTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCT 565
          | | || ||| |||| |||| || || | |||| | ||| | || | |||||||||
Db    421 CCCCTTATAAATCCTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCT 480

Qy    566 GGAAACCCTGAACACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATT 625
          ||| |||| | |||| ||||||||||| | || || || ||||||| ||||| |||
Db    481 GGAGACCCCAACTACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATT 540

Qy    626 CCTCTCTCTGTGATGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGC 685
          ||||| | ||||||||| ||||| || ||||||||| | ||||| ||||| |||||
Db    541 CCTCTTTTTGTGATGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAAT 600

Qy    686 CAGCAGCAAGCAACTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTT 745
          |||| || ||||| ||||| || || || || | |||| || |||| ||
Db    601 AGGCAGGTTGCTACTGCTCTGCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTG 660

Qy    746 GTGATCTTCTCTATACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCA 805
          || ||||||||| | || || ||||||||||| ||||||| ||| ||||||||| |||
Db    661 GTAATCTTCTCTGTGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCA 720

Qy    806 CGCCTGGATAGTTG---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACA 862
          ||||||| ||||| || | || || ||| | |||||| || | ||||
Db    721 CGCCTGGGGAGTTGGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATT 780

Qy    863 CTGACACGGCCTCTGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATG 922
          |||||||||| |||||||||||||||||||| |||||| || ||||||| || || ||
Db    781 GTGACACGGCCTTTGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTG 840

Qy    923 GGAGACCATTACAGAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACA 982
          ||||| || | ||| || ||||||||| | | | | ||||| ||||||| |||||||||
Db    841 GGAGATCACTTCAGGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACA 900

Qy    983 TCCTTCAGGACATGAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAATGAGACACTTGAT 1042
          ||||| || | ||| ||| || | | |||| |||| ||||||| | | |
Db    901 TCCTTTAGCAGATGGGCTCATGAACTCCTACTTTCATTCAGAGAAAATGATTCTCCTTCC 960

Qy   1043 AAACAGTGCTGTGCAGTTGAGTTT 1066
          || | || | || | ||
Db    961 TCACCCTCCTCAAATGGTGCGATT 984
```

RESULT 15
ADE85578/c
ID   ADE85578 standard; DNA; 639 BP.
XX
AC   ADE85578;
XX
DT   29-JAN-2004 (first entry)
XX

```
   Query Match              10.3%;   Score 158.8;  DB 9;  Length 639;
   Best Local Similarity    72.2%;   Pred. No. 6.8e-30;
   Matches  221;  Conservative     0;  Mismatches    82;  Indels    3;  Gaps    1;

Qy       727 CCTGGTGGTCCTGGCGGTTGTGATCTTCTCTATACTCTTCACACCCTATCATATCATGCG 786
                 | |||   || ||||  | || ||  ||||| | || || ||||| ||||| |||||||
Db       625 CTTGGTCATCATGGCACTGGTAATTATCTCTGTGCTATTAACACCATATCACGTCATGCG 566

Qy       787 CAATTTGAGGATCGCCTCACGCCTGGATAGTTG---GCCACAAGGATGTACACAGAAGGC 843
                 || |||||||||| ||||||||||  |||||   ||   ||   || || |||   |
Db       565 GTATGTGAGGATCGCTTCACGCCTGGTGAGTTGAAAGCAGTATCAGTGCACTCAGGTCGT 506

Qy       844 CATCAAATCTATATACACACTGACACGGCCTCTGGCCTTTCTGAACAGTGCCATCAATCC 903
                 |||||| ||  | ||||   ||||||||||| ||||||||||||||||| |||||| ||
```

```
Db    505 CATCAACTCCTTTTACATTGTGACACGGCCTTTGGCCTTTCTGAACAGTGTCATCAACCC 446

Qy    904 CATCTTCTACTTCCTCATGGGAGACCATTACAGAGAGATGCTGATTAGTAAGTTCAGACA 963
          |||||||  ||  ||  |||||||||  ||  |  |||  ||  |||||||||  |  |  |  |  |||||
Db    445 TGTCTTCTATTTTCTTATGGGAGATCACTTCAGGGACATGCTGATGAATCAACTGAGACA 386

Qy    964 ATACTTCAAGTCCCTTACATCCTTCAGGACATGAGCTGCTGGATGCAGGTCTTCACTCAG 1023
          |||||||  |||||||||||||||  ||  |  |||  |||   ||  |  |          ||||  ||||
Db    385 CAACTTCAAATCCCTTACATCCTTTAGCAGATGGGCTCATGAACTCCTACTTTCATTCAG 326

Qy   1024 CCAAAA 1029
          ||||
Db    325 AGAAAA 320


Search completed: August 24, 2004, 13:06:51
Job time : 661 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004   Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 12:29:30 ; Search time 124 Seconds
                                            (without alignments)
                                            6905.558 Million cell updates/sec

Title: US-09-891-138A-1
Perfect score: 1543
Sequence: 1 gctcctggcagagttttctg..........tgcctaaataaatcaatata 1543

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_NA:*
                 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
                 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
                 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
                 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
                 5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
                 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	589.2	38.2	1996	2	US-08-559-524A-1	Sequence 1, Appli
2	589.2	38.2	1996	3	US-08-749-707-1	Sequence 1, Appli
3	589.2	38.2	1996	4	US-09-947-922-1	Sequence 1, Appli
4	88.4	5.7	1429	4	US-09-016-434-1068	Sequence 1068, Ap
5	86.4	5.6	3055	4	US-09-016-434-1456	Sequence 1456, Ap
6	85.4	5.5	2025	4	US-09-016-434-1482	Sequence 1482, Ap
7	82.8	5.4	1571	4	US-09-016-434-1108	Sequence 1108, Ap
8	82.2	5.3	1805	4	US-08-405-271A-18	Sequence 18, Appl
9	82.2	5.3	1973	4	US-09-016-434-1391	Sequence 1391, Ap
10	82.2	5.3	1973	4	US-09-023-655-1417	Sequence 1417, Ap
11	82.2	5.3	3205	4	US-09-976-594-171	Sequence 171, App

12	80	5.2	1158	4	US-09-023-655-992	Sequence 992, App
13	80	5.2	1586	1	US-08-461-244-1	Sequence 1, Appli
14	80	5.2	1953	4	US-09-016-434-1096	Sequence 1096, Ap
15	80	5.2	2608	4	US-09-023-655-955	Sequence 955, App
16	79.6	5.2	984	3	US-08-513-974B-57	Sequence 57, Appl
17	79.6	5.2	984	4	US-09-461-436B-57	Sequence 57, Appl
18	79.6	5.2	1023	3	US-08-513-974B-379	Sequence 379, App
19	79.2	5.1	998	4	US-08-432-174A-3	Sequence 3, Appli
20	78.4	5.1	1495	4	US-09-016-434-1190	Sequence 1190, Ap
21	78.4	5.1	1495	4	US-09-023-655-1021	Sequence 1021, Ap
22	78.4	5.1	2156	1	US-08-012-988A-1	Sequence 1, Appli
23	78.4	5.1	2156	4	US-09-023-655-1247	Sequence 1247, Ap
24	78	5.1	1640	3	US-08-781-250-1	Sequence 1, Appli
25	77.6	5.0	1773	4	US-09-016-434-1405	Sequence 1405, Ap
26	76	4.9	1098	4	US-09-503-219B-7	Sequence 7, Appli
27	76	4.9	1255	4	US-09-023-655-993	Sequence 993, App
28	76	4.9	1679	4	US-09-016-434-1097	Sequence 1097, Ap
29	76	4.9	2100	4	US-09-495-050A-289	Sequence 289, App
30	74.8	4.8	984	3	US-08-459-046-1	Sequence 1, Appli
31	74.8	4.8	984	4	US-09-102-710B-1	Sequence 1, Appli
32	74.2	4.8	1140	4	US-09-016-434-750	Sequence 750, App
33	74.2	4.8	1301	2	US-08-467-948A-7	Sequence 7, Appli
34	74.2	4.8	1301	3	US-08-467-947A-7	Sequence 7, Appli
35	74	4.8	1086	4	US-09-170-496D-77	Sequence 77, Appl
36	74	4.8	1086	4	US-09-170-496D-205	Sequence 205, App
37	74	4.8	1643	1	US-08-383-750-3	Sequence 3, Appli
38	74	4.8	1643	3	US-08-352-678-3	Sequence 3, Appli
39	74	4.8	1643	4	US-09-536-954-3	Sequence 3, Appli
40	74	4.8	1643	4	US-09-016-434-1279	Sequence 1279, Ap
41	74	4.8	1643	5	PCT-US93-09636-3	Sequence 3, Appli
42	73.8	4.8	1065	3	US-08-847-296B-2	Sequence 2, Appli
43	73.8	4.8	1071	4	US-08-567-882-6	Sequence 6, Appli
44	73.8	4.8	1116	4	US-08-720-565-5	Sequence 5, Appli
45	73.8	4.8	1193	4	US-08-720-565-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-559-524A-1
; Sequence 1, Application US/08559524A
; Patent No. 5871963
; GENERAL INFORMATION:
;   APPLICANT: Conley, Pamela B.
;   APPLICANT: Jantzen, Hans-Michael
;   TITLE OF INVENTION: NOVEL PURINERGIC RECEPTOR
;   NUMBER OF SEQUENCES: 14
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
;   STREET: 1800 M Street, N.W.
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: USA
;   ZIP: 20036-5869
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk

```
;        COMPUTER:  IBM PC compatible
;        OPERATING SYSTEM:  PC-DOS/MS-DOS
;        SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;        APPLICATION NUMBER:  US/08/559,524A
;        FILING DATE:  15-NOV-1995
;        CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;        NAME:  Adler, Reid G.
;        REGISTRATION NUMBER:  30,988
;        REFERENCE/DOCKET NUMBER:  044481-5010-00-US
;     TELECOMMUNICATION INFORMATION:
;        TELEPHONE:  202-467-7000
;        TELEFAX:  202-467-7176
;  INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;        LENGTH:  1996 base pairs
;        TYPE:  nucleic acid
;        STRANDEDNESS:  single
;        TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;        NAME/KEY:  CDS
;        LOCATION:  625..1626
US-08-559-524A-1
```

```
  Query Match              38.2%;  Score 589.2;  DB 2;  Length 1996;
  Best Local Similarity    75.1%;  Pred. No. 1e-156;
  Matches  762;  Conservative   0;  Mismatches  248;  Indels   4;  Gaps    2;

Qy       39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
            | |   ||||||  ||||  | |||| | || ||| |||||  |||||||| | || | |
Db      632 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 691

Qy       99 AGTACTACCTCTCTGCATTTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
            |||||||||| ||     |||||||  || |||||| || | |||| | ||||| |||  ||
Db      692 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 751

Qy      159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
            ||| || | ||||||| |||||| |||||    |||||||||||||||||||| ||| | ||||| |
Db      752 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 811

Qy      219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
            |||||||| || ||||||||||| ||||| ||||||||||| ||||| ||||||| |||||
Db      812 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 871

Qy      279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
            ||||||||| ||   |||  |||||||| || ||||| ||||||||||||||||||||||||
Db      872 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 931

Qy      339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
            |  ||||||||||| ||||||||| |||||| ||||||||| || |||||| || ||||| ||
Db      932 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 991

Qy      399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
            | || ||||| |||||||||||||||||| |||| |||||||||| || ||||| ||||||||||
Db      992 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCT 1051
```

```
Qy    459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
          |  ||||  | ||||  ||||| |||||||||  | ||||||||| || |  | || |||
Db   1052 CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 1111

Qy    519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
          ||||    || || | |||| | ||| |  || | |||||||||||||||| |||| |
Db   1112 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 1171

Qy    579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
          ||||  ||||||||||| | || || || |||||||| ||||| ||||||||| |||||
Db   1172 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 1231

Qy    639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
          |||| |||||| || ||||||||| | ||||| ||| |  | ||||| |||||    || |
Db   1232 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 1291

Qy    699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
          |||| |||||| || || || || |     | |||| || |||| || || |||||||||
Db   1292 CTGCTCTGCCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 1351

Qy    759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
          | || || ||||||||||| | |||||| ||||||| ||||||||||| ||||||||| ||||
Db   1352 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 1411

Qy    819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
          |  ||  |     || || |||  | ||||||| | | |||    |||||||| ||
Db   1412 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGGCTT 1471

Qy    876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
          ||| ||||||||||||||| |||||| ||||||| || ||||||| || || |||||| || | ||
Db   1472 TGGGCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 1531

Qy    936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
          | || |||||||| | | |  | ||||| ||||||| |||||||||||||||| || | ||
Db   1532 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 1591

Qy    996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
          | ||| || | |    |||| ||||  |||| ||||    |||| ||||||
Db   1592 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAAGTGAGGGGCTTGTGAAACAG 1645
```

RESULT 2
US-08-749-707-1
; Sequence 1, Application US/08749707
; Patent No. 6063582
; GENERAL INFORMATION:
;   APPLICANT: Conley, Pamela B.
;   APPLICANT: Jantzen, Hans-Michael
;   TITLE OF INVENTION: NOVEL PURINERGIC RECEPTOR
;   NUMBER OF SEQUENCES: 14
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
;   STREET: 1800 M Street, N.W.
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: USA

; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1996 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 625..1626
US-08-749-707-1

  Query Match   38.2%;   Score 589.2;   DB 3;   Length 1996;
  Best Local Similarity   75.1%;   Pred. No. 1e-156;
  Matches 762; Conservative   0; Mismatches 248; Indels   4; Gaps   2;

Qy   39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
              | | ||||||| |||| | |||| | || ||| ||||| ||||||||| | || | |
Db   632 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 691

Qy   99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
              ||||||||||| || |||||||| || ||||||| || | ||| | ||||| ||| ||
Db   692 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 751

Qy   159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
              ||| || | |||||||| |||||| |||||||||||||||||||||| ||| | |||||| |
Db   752 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 811

Qy   219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
              ||||||| || |||||||||| ||||| |||||||||||| ||||| ||||||| | |||||
Db   812 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 871

Qy   279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
              ||||||||| || ||| ||||||||| || ||||| |||||||||||||||||||||||||||||
Db   872 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 931

Qy   339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
              | ||||||||||| |||||||| ||||| ||||||||| || |||||| || ||||| | ||
Db   932 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 991

```
Qy     399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
           |  ||  |||||  |||||||||||||||||  ||||  |||||||||  ||  |||||  ||||||||||
Db     992 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAGAAAGAGTTTGCTATTTTAATCT 1051

Qy     459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
           |  ||||  |  |||| |||||  ||||||||  | |||||||||  ||  |  |  ||  |||
Db    1052 CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 1111

Qy     519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
           ||||    ||  ||  | ||||  | |||  |  ||  | ||||||||||||||  ||||  |
Db    1112 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 1171

Qy     579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
           ||||  |||||||||||  |  | ||  || ||  |||||||  ||||||  |||||||| |  |||||
Db    1172 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 1231

Qy     639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
           ||||  |||||  ||  ||||||||| |     |||||  |||||  |||||     ||||   || |
Db    1232 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 1291

Qy     699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
           ||||  |||||  ||  ||  ||  ||  |      | ||||  ||  ||||  ||  ||  ||||||||||
Db    1292 CTGCTCTGCCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 1351

Qy     759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
           |  ||  ||  ||||||||||||||||   |||||||  |||  ||||||||||| |||||||||||   ||||
Db    1352 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 1411

Qy     819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
           |    ||   |        ||  ||  |||   | ||||||| ||  | ||||      |||||||| ||
Db    1412 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGGCTT 1471

Qy     876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
           ||| |||||||||||||||  |||||| ||  |||||| ||  ||  |||||||  || | ||
Db    1472 TGGGCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 1531

Qy     936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
           | ||  |||||||| | | |  |  |||||  |||||||  |||||||||||||||||  || | ||
Db    1532 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 1591

Qy     996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
            | |||  ||  |    |    ||||  |||| |||| ||||  |||| ||||  ||||||
Db    1592 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAAGTGAGGGGCTTGTGAAACAG 1645
```

RESULT 3
US-09-947-922-1
; Sequence 1, Application US/09947922
; Patent No. 6680373
;   GENERAL INFORMATION:
;   APPLICANT: Conley, Pamela B.
;   Jantzen, Hans-Michael
;   TITLE OF INVENTION: NOVEL PURINERGIC RECEPTOR
;   NUMBER OF SEQUENCES: 14
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
;   STREET: 1800 M Street, N.W.

```
;               CITY: Washington
;               STATE: D.C.
;               COUNTRY: USA
;               ZIP: 20036-5869
;        COMPUTER READABLE FORM:
;               MEDIUM TYPE: Floppy disk
;               COMPUTER: IBM PC compatible
;               OPERATING SYSTEM: PC-DOS/MS-DOS
;               SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;               APPLICATION NUMBER: US/09/947,922
;               FILING DATE: 07-Sep-2001
;               CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;               APPLICATION NUMBER: US/08/749,707
;               FILING DATE: 15-NOV-1996
;        ATTORNEY/AGENT INFORMATION:
;               NAME: Adler, Reid G.
;               REGISTRATION NUMBER: 30,988
;               REFERENCE/DOCKET NUMBER: 044481-5010-01-US
;        TELECOMMUNICATION INFORMATION:
;               TELEPHONE: 202-467-7000
;               TELEFAX: 202-467-7176
;    INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;               LENGTH: 1996 base pairs
;               TYPE: nucleic acid
;               STRANDEDNESS: single
;               TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        FEATURE:
;               NAME/KEY:  CDS
;               LOCATION:  625..1626
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-947-922-1

   Query Match            38.2%;   Score 589.2;  DB 4;   Length 1996;
   Best Local Similarity  75.1%;   Pred. No. 1e-156;
   Matches  762; Conservative   0;  Mismatches  248;  Indels   4;  Gaps    2;

Qy       39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
            | |   |||||| |||| | |||| | || ||| ||||| |||||||| | || | |
Db      632 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 691

Qy       99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
            |||||||||| ||     ||||||| || |||||| || | ||| | ||||| ||| ||
Db      692 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 751

Qy      159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
            ||| || | ||||||| |||||| |||||||||||||||||||| ||| | |||||| |
Db      752 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 811

Qy      219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
            ||||||| || ||||||||| ||||||| ||||||||||| ||||| ||||||| |||||
Db      812 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 871

Qy      279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
```

```
                   |||||||||  ||  |||  ||||||||  ||  |||||  ||||||||||||||||||||||||||
Db      872 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 931

Qy      339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
                |  |||||||||||  |||||||||  |||||  |||||||||  ||  ||||||  ||  ||||||  ||
Db      932 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 991

Qy      399 TCATGAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
                |  ||  ||||||  |||||||||||||||||  ||||  ||||||||||  ||  |||||  ||||||||||
Db      992 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAGAAAGAGTTTGCTATTTTAATCT 1051

Qy      459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
                |   ||||   |  ||||   |||||  |||||||||||  |  ||||||||||  ||  |  |  ||  |||
Db     1052 CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 1111

Qy      519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
                ||||     ||   ||  |  ||||  |  ||||  |   ||  | |||||||||||||||  |||||  |
Db     1112 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 1171

Qy      579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
                |||| ||||||||||||||  |  ||  ||  ||  ||||||||  ||||||  ||||||||  |  |||||
Db     1172 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTGTGA 1231

Qy      639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
                ||||  ||||||  ||  |||||||||  |      |||||  |||||  |||||      ||||    ||  |
Db     1232 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 1291

Qy      699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
                ||||  ||||||  ||  ||  ||  ||  |       |  ||||  ||  ||||  ||  ||  ||||||||||
Db     1292 CTGCTCTGCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 1351

Qy      759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
                |  ||  ||  |||||||||||  |||||||  |||  ||||||||||  ||||||||||     ||||
Db     1352 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 1411

Qy      819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
                |   ||  |    ||  ||  |||  |  ||||||| ||  |  ||||     |||||||||  ||
Db     1412 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGGCTT 1471

Qy      876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
                |||  ||||||||||||||||  |||||  ||  |||||||  ||  ||  |||||||  ||  |  ||
Db     1472 TGGGCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 1531

Qy      936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
                |  ||  |||||||||  |  |  |  |  |||||  |||||||  ||||||||||||||||||  ||  |  ||
Db     1532 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 1591

Qy      996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
                |  |||  ||  |   |     ||||  ||||  ||||  ||||     |||||  ||||||
Db     1592 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAGTGAGGGGCTTGTGAAACAG 1645
```

RESULT 4
US-09-016-434-1068
; Sequence 1068, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:

```
    Query Match              5.7%;   Score 88.4;  DB 4;  Length 1429;
    Best Local Similarity   45.7%;   Pred. No. 5.8e-15;
    Matches  385;  Conservative   0;  Mismatches  451;  Indels   6;  Gaps   2;

Qy       107 CTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGTG 166
             ||   |||      ||||||| | |  ||  |  | || ||       || | | |         |
Db       292 CTGCCTGTGAGCTATGCAGTTGTCTTTGTGCTGGGCTTGGGCCTTAACGCCCCAACCCTA 351

Qy       167 TTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCTT 226
             |    || | ||||| ||| |        |||| |    |        ||| | || ||||
Db       352 TGGCTCTTCATCTTCCGCCTCCGACCCTGGGATGCAACGGCCACCTACATGTTCCACCTG 411

Qy       227 TCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTAT---GCC 283
             |  | || |||       |    ||    | || |||| ||| || |    |||| |||
```

```
Db      412 GCATTGTCAGACACCTTGTATGTGCTGTCGCTGCCCACCCTCATCTACTATTATGCAGCC 471

Qy      284 AATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACC 343
                |   |   |  || ||| |||     ||    ||||  |     |  |||  | |  |  |
Db      472 CACAACCACTGGCCCTTTGGCACTGAGATCTGCAAGTTCGTCCGCTTTCTTTTCTATTGG 531

Qy      344 AACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGCTCATG 403
                |||||||||  |||    |||| ||||||||| |  ||| |||  | ||||  || |||  |||
Db      532 AACCTCTACTGCAGTGTCCTTTTCCTCACCTGCATCAGCGTGCACCGCTACCTGGGCATC 591

Qy      404 AAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCTCGCTG 463
                ||||   |  ||  |  ||                  |  | ||      |   ||| |||  |||
Db      592 TGCCACCCACTTCGGGCACTACGCTGGGGCCGCCCTCGCCTCGCAGGCCTTCTCTGCCTG 651

Qy      464 GCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATTCTGTC 523
                || || |||    |  | ||  ||        |  | |||| ||    || |  ||| |   |
Db      652 GCAGTTTGGTTGGTCGTAGCCGGCTGCCTCGTGCCCAACCTGTTCTTTGTCACAACCAGC 711

Qy      524 CCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAACACAAT 583
                ||||        |     ||| |||    | |   ||| | |   ||| ||| ||
Db      712 AACAAAGGGACCACCGTCCTGTGCCATGACACCACTCGGCCTGAAGAGTTTGACCACTAT 771

Qy      584 CTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGATGTGC 643
                |    |  | ||||      |  |    | || |||     | ||    | ||  |  || |
Db      772 GTGCACTTCAGCTCGGCGGTCATGGGGCTGCTCTTTGGCGTGCCCTGCCTGGTCACTCTT 831

Qy      644 TTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAACTGCC 703
                | |  |||       |||| |     |  |    |    |        |  |||    |
Db      832 GTTTGCTATGGACTCATGGCTCGTCGCCTGTATCAGCCCTTGCCAGGCTCTGCACAGTCG 891

Qy      704 CTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTATACTC 763
                |  | |     |   | |  ||  | |   ||| || |    |||| ||| |   |
Db      892 TCTTCTCGCCTCCGCTCTCTCCGCACCATAGCTGTGGTGCTGACTGTCTTTGCTGTCTGC 951

Qy      764 TTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTTGGCCA 823
                ||| || | |  || ||| |||| |||| |     |  |  | |  || | |||       |
Db      952 TTCGTGCCTTTCCACATCACCCGCACCATTTACTACCTGGCCAGGCTGTTGGAA---GCT 1008

Qy      824 CAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTCTGGCCTTT 883
                |   |   |   || |||    ||||    | || | | |||| ||||| |||||| ||||||  |
Db     1009 GACTGCCGAGTACTGAACATTGTCAACGTGGTCTATAAAGTGACTCGGCCCCTGGCCAGT 1068

Qy      884 CTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACAGAGAGATG 943
                |||||  | |  |||| |  |||||||| ||||  || ||| | || | ||  ||       |
Db     1069 GCCAACAGCTGCCTGGATCCTGTGCTCTACTTGCTCACTGGGGACAAATATCGACGTCAG 1128

Qy      944 CT 945
                ||
Db     1129 CT 1130
```

RESULT 5
US-09-016-434-1456
; Sequence 1456, Application US/09016434
; Patent No. 6500938
;   GENERAL INFORMATION:
;   APPLICANT: Janice Au-Young

```
    Query Match            5.6%;  Score 86.4;  DB 4;  Length 3055;
    Best Local Similarity  46.1%;  Pred. No. 3.4e-14;
    Matches  402;  Conservative    0;  Mismatches  461;  Indels   9;  Gaps    3;

Qy       80 ACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTT 139
            ||  || |   ||| |    ||||||||| | ||  | ||    | |  |||||  |
Db      982 ACCAAGACGGGCTTCCAGTTTTACTACCTGCCGGCTGTCTACATCTTGGTATTCATCATC 1041

Qy      140 GGACTGCTTGGGAATGTCACTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAAC 199
            ||  | || || || |    |    || |     | | |||||  |||||||  ||||| |
Db     1042 GGCTTCCTGGGCAACAGCGTGGCCATCTGGATGTTCGTCTTCCACATGAAGCCCTGGAGC 1101

Qy      200 AGCAGCAATGTCTATCTTTTTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTT 259
            |||  |   || || ||  | ||| ||  | | | |||||| | ||  | || || ||
Db     1102 GGCATCTCCGTGTACATGTTCAATTTGGCTCTGGCCGACTTCTTGTACGTGCTGACTCTG 1161
```

```
Qy      260  CCCATCCTGATAAAGAGTTATGCCAATGATA---AGGGGACCTATGGAGATGTTCTCTGT  316
             ||   ||||||          ||   |||| | |    |   ||| || || ||||   | |||
Db     1162  CCAGCCCTGATCTTCTACTACTTCAATAAAACAGACTGGATCTTCGGGGATGCCATGTGT  1221

Qy      317  ATAAGCAACCGATATGTGCTTCACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTC  376
             | |   | | | |   |  ||||   ||||||||   |||| | | || || || |
Db     1222  AAACTGCAGAGGTTCATCTTTCATGTGAACCTCTA---TGGCATCTTGTTTCTGACATGC  1278

Qy      377  ATTAGCATGGACCGATATCTGCTCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAG  436
             || ||     |||| ||        ||  ||||||  ||      |    |   |   | ||
Db     1279  ATCAGTGCCCACCGGTACAGCGGTGTGGTGTACCCCCTCAAGTCCCTGGGCCGGCTCAAA  1338

Qy      437  AAGGAATTTGCCATTTTAATCTCGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTA  496
             ||| |   ||| || |   |||      ||   || ||| | | |||      | |    |
Db     1339  AAGAAGAATGCGATCTGTATCAGCGTGCTGGTGTGGCTCATTGTGGTGGTGGCGATCTCC  1398

Qy      497  CCCATGCTCACTTTCATCAATTCT---GTCCCAAAAGAAGAGGGCAGTAACTGCATCGAC  553
             ||||| |||   || |     | |    ||||   ||| |  |   || | |||    ||||
Db     1399  CCCATCCTCTTCTACTCAGGTACCGGGGTCCGCAAAAACAAAACCATCACCTGTTACGAC  1458

Qy      554  TATGCAAGTTCTGGAAACCCTGAACACAATCTCATTTACAGCCTCTGCCTGACTTTGTTG  613
             |        |   |||   |       || |||| |||||| |||| | ||| |||   ||
Db     1459  ACCACCTCAGACGAGTACCTGCGAAGTTATTTCATCTACAGCATGTGCACGACCGTGGCC  1518

Qy      614  GGCTTCCTAATTCCTCTCTCTGTGATGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTA  673
             |||    | ||  |       ||||      || |||      || ||          ||
Db     1519  ATGTTCTGTGTCCCCCTTGGTGCTGATTCTGGGCTGTTACGGATTAATTGTGAGAGCTTTG  1578

Qy      674  AAGAGGAGGAGCCAGCAGCAAGCAACTGCCCTGCCACTGGACAAACCCCAACGCCTGGTG  733
             |      |     | | | | | |  ||     | |       |    ||       | |
Db     1579  ATTTACAAAGATCTGGACAACTCTCCTCTGAGGAGAAAATCGATTTACCTGGTAATCATT  1638

Qy      734  GTCCTGGCGGTTGTGATCTTCTCTATACTCTTCACACCCTATCATATCATGCGCAATTTG  793
             || ||| | ||| |      | ||| ||        |       | | |    || |||
Db     1639  GTACTGACTGTTTTTGCTGTGTCTTACATCCCTTTCCATGTGATGAAAACGATGAACTTG  1698

Qy      794  AGGATCGCCTCACGCCTGGATAGTTGGCCACAAGGATGTACACAGAAGGCCATCAAATCT  853
             |||   |            |   | |     ||    | |   |   | |      | |
Db     1699  AGGGCCCGGCTTGATTTTCAGACCCCAGCAATGTGTGCTTTCAATGACAGGGTTTATGCC  1758

Qy      854  ATATACACTGACACGGCCTCTGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTAC  913
             | ||      |||||| |   ||| ||   ||| |||||||     |   |||||| ||||
Db     1759  ACGTATCAGGTGACAAGAGGTCTAGCAAGTCTCAACAGTTGTGTGGACCCCATTCTCTAT  1818

Qy      914  TTCCTCATGGGAGACCATTACAGAGAGATGCT  945
             ||| |    ||||||  || ||||  ||  ||
Db     1819  TTCTTGGCGGGAGATACTTTCAGAAGGAGACT  1850
```

RESULT 6
US-09-016-434-1482
; Sequence 1482, Application US/09016434
; Patent No. 6500938
;   GENERAL INFORMATION:
;   APPLICANT: Janice Au-Young
;   APPLICANT: Jeffrey J. Seilhamer

```
;     TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF SIGNALING
;     TITLE OF INVENTION:  PATHWAY GENE EXPRESSION
;     NUMBER OF SEQUENCES: 1490
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;       STREET:  3174 PORTER DRIVE
;       CITY:  PALO ALTO
;       STATE:  CALIFORNIA
;       COUNTRY:  USA
;       ZIP:  94304
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/016,434
;       FILING DATE:  HEREWITH
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Zeller, Karen J.
;       REGISTRATION NUMBER:  37,071
;       REFERENCE/DOCKET NUMBER:  PA-0002 US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (650) 855-0555
;       TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  1482:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2025 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     IMMEDIATE SOURCE:
;       LIBRARY:  GENBANK
;       CLONE:  g984506
US-09-016-434-1482

   Query Match               5.5%;  Score 85.4;  DB 4;  Length 2025;
   Best Local Similarity    46.5%;  Pred. No. 5.1e-14;
   Matches  389;  Conservative   0;  Mismatches  436;  Indels  12;  Gaps   3;

Qy         91 CTTGAATAAGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGG 150
              ||| || |       || |  ||| | ||| |  | | || | | |||| ||| |
Db        335 CTTCAAGTACGTGCTGCTGCCTGTGTCCTACGGCGTGGTGTGCGTGCTTGGGCTGTGTCT 394

Qy        151 GAATGTCACTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGT 210
              |||| | |   | | | | |  || | | | | | || | |||| ||||||      ||
Db        395 GAACGCCGTGGCGCTCTACATCTTCTTGTGCCGCCTCAAGACCTGGAATGCGTCCACCAC 454

Qy        211 CTATCTTTTTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGAT 270
              ||| | ||  |||| |  | | ||||        |    | | |||| || | ||| |
Db        455 ATATATGTTCCACCTGGCTGTGTCTGATGCACTGTATGCGGCCTCCCTGCCGCTGCTGGT 514
```

```
Qy      271  AAAGAGTTATGCCAAT---GATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCG 327
                |    || |||      || |  || |||    |   || ||||| |        ||
Db      515  CTATTACTACGCCCGCGGCGACCACTGGCCCTTCAGCACGGTGCTCTGCAAGCTGGTGCG 574

Qy      328  ATATGTGCTTCACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGA 387
                |   |   |   ||||||||||| |||   ||||||||||||||||||| | ||| ||| || |
Db      575  CTTCCTCTTCTACACCAACCTTTACTGCAGCATCCTCTTCCTCACCTGCATCAGCGTGCA 634

Qy      388  CCGATATCTGCTCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGC 447
                ||| | |||| | |          ||| | ||     |                |      | ||
Db      635  CCGGTGTCTGGGCGTCTTACGACCTCTGCGCTCCCTGCGCTGGGGCCGGGCCCGCTACGC 694

Qy      448  CATTTTAATCTCGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCAC 507
                   |  |   ||| || ||||  || |||              ||| |||||
Db      695  TCGCCGGGTGGCCGGGGCCGTGTGGGTGTTGGTGCTGGCCTGCCAGGCCCCCGTGCTCTA 754

Qy      508  TTTCATCAATTCTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGG 567
                ||  |||  |  |  | |      | |   ||  | ||||  ||||       |     | |
Db      755  CTTTGTCACCACCAGCGCGCGCGGGGGCCGCGTAACCTGCCACGACACCTCGGCACCCGA 814

Qy      568  AAACCCTGAACACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCC 627
                   |      | |     |      ||||||     | ||       || ||||    | ||
Db      815  GCTCTTCAGCCGCTTCGTGGCCTACAGCTCAGTCATGCTGGGCCTGCTCTTCGCGGTGCC 874

Qy      628  TCTCTCTGTGATGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCA 687
                |  | || ||       ||| |||        ||||          |   | |      | |
Db      875  CTTTGCCGTCATCCTTGTCTGTTACGTGCTCATGGCTCGGCGACTGCTAAAGCCAGCCTA 934

Qy      688  GCAGCAAGCAACTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGT 747
                |    |    | ||| ||  || |||  | |    | ||       | | ||  || || ||
Db      935  CGGGACCTCGGGCGGCCTCCCTAGGGCCAAGCGCAAGTCCGTGCGCACCATCGCCGTGGT 994

Qy      748  G------ATCTTCTCTATACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGC 801
                |       |||||| |  |   |||| ||  || || ||| |||| |       | |
Db      995  GCTGGCTGTCTTCGCCCTCTGCTTCCTGCCATTCCACGTCACCCGCACCCTCTACTACTC 1054

Qy      802  CTCACGCCTGGATAGTTGGCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACAC 861
                ||  |||  |         |      |      || |  || |||||||||       |||| 
Db     1055  CTTCCGCTCGCTGG---ACCTCAGCTGCCACACCCTCAACGCCATCAACATGGCCTACAA 1111

Qy      862  ACTGACACGGCCTCTGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCT 918
                | || ||||||| |||||| |    |||||| | |  | ||| |  ||||||||||||
Db     1112  GGTTACCCGGCCGCTGGCCAGTGCTAACAGTTGCCTTGACCCCGTGCTCTACTTCCT 1168
```

RESULT 7
US-09-016-434-1108
; Sequence 1108, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

```
;        STREET:  3174 PORTER DRIVE
;        CITY:  PALO ALTO
;        STATE:  CALIFORNIA
;        COUNTRY:  USA
;        ZIP:  94304
;     COMPUTER READABLE FORM:
;        MEDIUM TYPE:  Floppy disk
;        COMPUTER:  IBM PC compatible
;        OPERATING SYSTEM:  PC-DOS/MS-DOS
;        SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;     CURRENT APPLICATION DATA:
;        APPLICATION NUMBER:  US/09/016,434
;        FILING DATE:  HEREWITH
;        CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;        APPLICATION NUMBER:
;        FILING DATE:
;        CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;        NAME:  Zeller, Karen J.
;        REGISTRATION NUMBER:  37,071
;        REFERENCE/DOCKET NUMBER:  PA-0002 US
;     TELECOMMUNICATION INFORMATION:
;        TELEPHONE:  (650) 855-0555
;        TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  1108:
;     SEQUENCE CHARACTERISTICS:
;        LENGTH:  1571 base pairs
;        TYPE:  nucleic acid
;        STRANDEDNESS:  single
;        TOPOLOGY:  linear
;     IMMEDIATE SOURCE:
;        LIBRARY:  GENBANK
;        CLONE:  g1296659
US-09-016-434-1108

   Query Match              5.4%;   Score 82.8;   DB 4;   Length 1571;
   Best Local Similarity   46.2%;   Pred. No. 2.4e-13;
   Matches  390;  Conservative    0;  Mismatches  442;  Indels   12;  Gaps     3;

Qy        89 ATCTTGAATAAGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTT 148
             | ||| ||   |          ||  |  |  | ||| |    | | |       ||| ||||
Db       343 AACTTCAAGCAACTGCTGCTGCCACCTGTGTATTCGGCGGTGCTGGCGGCTGGCCTGCCG 402


Qy       149 GGGAATGTCACTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAAT 208
             ||| ||  ||| |    |        |  |  | |        |  |  || |  ||| 
Db       403 CTGAACATCTGTGTCATTACCCAGATCTGCACGTCCCGCCGGGCCCTGACCCGCACGGCC 462


Qy       209 GTCTATCTTTTTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTG 268
             || ||      | ||||||  | |  |||||| |   |       ||| |||| ||| | ||
Db       463 GTGTACACCCTAAACCTTGCTCTGGCTGACCTGCTATATGCCTGCTCCCTGCCCCTGCTC 522


Qy       269 ATAAAGAGTTATGCCAA---TGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAAC 325
             ||   | | ||||||| |    |||| |  || ||| ||| || |   |||          |
Db       523 ATCTACAACTATGCCCAAGGTGATCACTGGCCCTTTGGCGACTTCGCCTGCCGCCTGGTC 582


Qy       326 CGATATGTGCTTCACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATG 385
```

```
                   ||  |    |    |    |   ||||||| ||  ||||||||||||||||||||| | ||| ||| |
Db     583 CGCTTCCTCTTCTATGCCAACCTGCACGGCAGCATCCTCTTCCTCACCTGCATCAGCTTC 642

Qy     386 GACCGATATCTGCTCATGAAGTACCCTTT---CCGAGAACACTTTCTACAAAAGAAGGAA 442
                | || || |||  |||      |||| |   ||     ||| |
Db     643 CAGCGCTACCTGGGCATCTGCCACCCGCTGGCCCCCTGGCACAAACGTGGGGGCCGCCGG 702

Qy     443 TTTGCCATTTTAATCTCGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATG 502
                ||||     || | |    | || || |||     ||||| |    || ||||
Db     703 GCTGCCTGGCTAGTGTGTGTAGCCGTGTGGCTGGCCGTGACAACCCAGTGCCTGCCCACA 762

Qy     503 CTCACTTTCATCAATTCTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGT 562
                || |||      | |  |     |     | ||| | ||||   |||
Db     763 GCCATCTTCGCTGCCACAGGCATCCAGCGTAACCGCACTGTCTGCTATGACCTCAGCCCG 822

Qy     563 TCTGGAAACCCTGAACACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTA 622
                |||     |   ||| ||||    || || |    || ||| | | |||||||||
Db     823 CCTGCCCTGGCCACCCACTATATGCCCTATGGCATGGCTCTCACTGTCATCGGCTTCCTG 882

Qy     623 ATTCCTCTCTGTGATGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGG 682
                | ||  |  |||   ||     || ||||| |    ||| |  |  | |    | |   |
Db     883 CTGCCCTTTGCTGCCCTGCTGGCCTGCTACTGTCTCCTGGCCTGCCGCCTGTGCCGCCAG 942

Qy     683 AGCCAGCAGCAAGCAACTG------CCCTGCCACTGGACAAACCCCAACGCCTGGTGGTC 736
                | | || | |||       || | | ||| |   ||| ||| ||
Db     943 GATGGCCCGGCAGAGCCTGTGGCCCAGGAGCGGCGTGGCAAGGCGGCCCGCATGGCCGTG 1002

Qy     737 CTGGCGGTTGTGATCTTCTCTATACTCTTCACACCCTATCATATCATGCGCAATTTGAGG 796
                ||| || ||    ||| | ||    |||| || | ||| ||||     |
Db    1003 GTGGTGGCTGCTGCCTTTGCCATCAGCTTCCTGCCTTTTCACATCACCAAGACAGCCTAC 1062

Qy     797 ATCGCCTCACGCCTGGATAGTTGGCCACAAGGATGTACACAGAAGGCCATCAAATCTATA 856
                | ||   |||  |      |     |   |   |   | |   | ||||| |    | |
Db    1063 CTGGCAGTGCGCTCGACGCCGGGCGTCCCCTGCACTGTATTGGAGGCCTTTGCAGCGGCC 1122

Qy     857 TACACACTGACACGGCCTCTGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTC 916
                |||| |    ||  ||||||  | |||   |     ||||| |    | |||||| ||| || |
Db    1123 TACAAAGGCACGCGGCCGTTTGCCAGTGCCAACAGCGTGCTGGACCCCATCCTCTTCTAC 1182

Qy     917 CTCA 920
                |||
Db    1183 TTCA 1186
```

RESULT 8
US-08-405-271A-18
; Sequence 18, Application US/08405271A
; Patent No. 6432652
;   GENERAL INFORMATION:
;   APPLICANT: EVANS, CHRISTOPHER J.
;   APPLICANT: KEITH, DUANE E.
;   TITLE OF INVENTION: OPIOID RECEPTOR GENES
;   NUMBER OF SEQUENCES: 25
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: MORRISON & FOERSTER
;   STREET: 2000 PENNSYLVANIA AVENUE, NW, Suite 5500
;   CITY: WASHINGTON

```
   Query Match            5.3%;   Score 82.2;  DB 4;  Length 1805;
   Best Local Similarity   44.5%;  Pred. No. 3.8e-13;
   Matches  379;  Conservative   0;  Mismatches  463;  Indels   9;  Gaps   1;


Qy        85 GGCTATCTTGAATAAGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACT 144
             | |  || |   ||| |  | ||  |  | ||    || ||  |    || ||
Db       147 GCCCCTCGGGCTCAAGGTCACCATCGTGGGGCTCTACCTGGCCGTGTGTGTCGGAGGGCT 206


Qy       145 GCTTGGGAATGTCACTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAG 204
             || ||||||   |  |||  ||| || |  ||||     || |||  |     ||    ||
Db       207 CCTGGGGAACTGCCTTGTCATGTACGTCATCCTCAGGCACACCAAAATGAAGACAGCCAC 266


Qy       205 CAATGTCTATCTTTTTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCAT 264
             |||| | ||  | ||||||||| || |  | ||| |     ||||| || || ||| |
Db       267 CAATATTTACATCTTTAACCTGGCCCTGGCCGACACTCTGGTCCTGCTGACGCTGCCCTT 326


Qy       265 CCTGATAAAGAGTTATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAA 324
             || |   | |         |   |     || | | ||| ||| || || | |
Db       327 CCAGGGCACGGACATCCTCCTGGGCTTCTGGCCGTTTGGGAATGCGCTGTGCAAGACAGT 386


Qy       325 CCGATATGTGCTTCACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCAT 384
             |        |  || |||| | | |||||||| | || ||| ||| ||| || |
Db       387 CATTGCCATTGACTACTACAACATGTTCACCAGCACCTTCACCCTAACTGCCATGAGTGT 446


Qy       385 GGACCGATATCTGCTCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATT 444
             ||| || ||| |   |||    |||| |||| | |   | |  | |   |
```

```
Db    447 GGATCGCTATGTAGCCATCTGCCACCCCATCCGTGCCCTCGACGTCCGCACGTCCAGCAA 506

Qy    445 TGCCATTTTAATCTCGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCT 504
          |||        ||    |||| |||||||| || |  | | |  | | | ||| |
Db    507 AGCCCAGGCTGTCAATGTGGCCATCTGGGCCCTGGCCTCTGTTGTCGGTGTTCCCGTTGC 566

Qy    505 CACTTTCATCAATTCTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTC 564
          ||  |   |    |       | || || |  | | ||| | |   |        | |
Db    567 CATCATGGGCTCGGCACAGGTCGAGGATGAAGAGATCGAGTGCCTGGTGGAGATCCCTAC 626

Qy    565 TGGAAACCCTGAACACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAAT 624
            |  | ||        |   | |   | ||||| |      | ||   |||| |  |
Db    627 CCCTCAGGATTACTGGGGCCCGGTGTTTGCCATCTGCATCTTCCTCTTCTCCTTCATCGT 686

Qy    625 TCCTCTCTGTGATGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAG 684
          || |   || || |   ||| |||| |   ||| |        |      |      |
Db    687 CCCCGTGCTCGTCATCTCTGTCTGCTACAGCCTCATGATCCGGCGGCTCCGTGGAGTCCG 746

Qy    685 CCAGCAGCAAGCAACTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGT 744
          || ||      |    |         ||| |  | ||| |       |||| | |
Db    747 CCTGCTCTCGGGCTCCCGAGAGAAGGACCGGAACCTGCGGCGCATCACTCGGCTGGTGCT 806

Qy    745 TGTGATCTTCTCTATACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTC 804
          |||  |   |   || |  || |         ||      | |    ||   | | |||
Db    807 GGTGGTAGTGGCTGTGTTCGTGGGCTGCTGGACGCCTGTCCAGGTCTTCGTGCTGGCCCA 866

Qy    805 ACGCCTGGATAGTTGGCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACT 864
          | | ||||    |   |||     |   ||     ||||||     | | |||
Db    867 AGGGCTGGGGGTTCAGCCGAGCAGCGAGACTGCCGTGGCCATTCTGCGCTTCTGCAC--- 923

Qy    865 GACACGGCCTCTGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGG 924
           | | |||| ||  | | ||||   | |||| | |||| ||||||| |||||  | ||  |||
Db    924 ------GGCCCTGGGCTACGTCAACAGCTGCCTCAACCCCATCCTCTACGCCTTCCTGGA 977

Qy    925 AGACCATTACA 935
          ||   | | ||
Db    978 TGAGAACTTCA 988
```

RESULT 9
US-09-016-434-1391
; Sequence 1391, Application US/09016434
; Patent No. 6500938
;   GENERAL INFORMATION:
;   APPLICANT: Janice Au-Young
;   APPLICANT: Jeffrey J. Seilhamer
;   TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
;   TITLE OF INVENTION: PATHWAY GENE EXPRESSION
;   NUMBER OF SEQUENCES: 1490
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 PORTER DRIVE
;   CITY: PALO ALTO
;   STATE: CALIFORNIA
;   COUNTRY: USA
;   ZIP: 94304
;   COMPUTER READABLE FORM:

```
Query Match              5.3%;   Score 82.2;   DB 4;   Length 1973;
Best Local Similarity   44.5%;   Pred. No. 4e-13;
Matches   379;  Conservative    0;  Mismatches  463;  Indels    9;  Gaps    1;


Qy        85 GGCTATCTTGAATAAGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACT 144
             | |  || |   ||| | | || |  | || |      || || |   || ||
Db       315 GCCCCTCGGGCTCAAGGTCACCATCGTGGGGCTCTACCTGGCCGTGTGTGTCGGAGGGCT 374


Qy       145 GCTTGGGAATGTCACTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAG 204
             || |||||  |   ||| ||| || |  ||||     || || |    ||      ||
Db       375 CCTGGGGAACTGCCTTGTCATGTACGTCATCCTCAGGCACACCAAAATGAAGACAGCCAC 434


Qy       205 CAATGTCTATCTTTTTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCAT 264
             |||| | || | ||||||||| || | | ||| |     ||||| || || ||| |
Db       435 CAATATTTACATCTTTAACCTGGCCCTGGCCGACACTCTGGTCCTGCTGACGCTGCCCTT 494


Qy       265 CCTGATAAAGAGTTATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAA 324
             || |   || |          |  | ||| | ||| ||| || || | |
Db       495 CCAGGGCACGGACATCCTCCTGGGCTTCTGGCCGTTTGGGAATGCGCTGTGCAAGACAGT 554


Qy       325 CCGATATGTGCTTCACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCAT 384
             |        |    ||  |||| | | |||||||| | ||  ||| |||  ||| || |
Db       555 CATTGCCATTGACTACTACAACATGTTCACCAGCACCTTCACCCTAACTGCCATGAGTGT 614


Qy       385 GGACCGATATCTGCTCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATT 444
             ||| || ||| |   |||     |||| |||| | | |   | | |  |
Db       615 GGATCGCTATGTAGCCATCTGCCACCCCATCCGTGCCCTCGACGTCCGCACGTCCAGCAA 674
```

```
Qy      445  TGCCATTTTAATCTCGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCT  504
             |||         ||      ||||  ||||||||| | |     | | |   | | | ||| |
Db      675  AGCCCAGGCTGTCAATGTGGCCATCTGGGCCCTGGCCTCTGTTGTCGGTGTTCCCGTTGC  734

Qy      505  CACTTTCATCAATTCTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTC  564
             ||   |   |    |        | || || | |   | ||| | |    |       | |
Db      735  CATCATGGGCTCGGCACAGGTCGAGGATGAAGAGATCGAGTGCCTGGTGGAGATCCCTAC  794

Qy      565  TGGAAACCCTGAACACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAAT  624
               |   | |         |   | |     | |||||| |     | ||    |||| | | |
Db      795  CCCTCAGGATTACTGGGGCCCGGTGTTTGCCATCTGCATCTTCCTCTTCTCCTTCATCGT  854

Qy      625  TCCTCTCTCTGTGATGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAG  684
             ||  |    || || |   ||| ||||  |   ||| |        |      |      |
Db      855  CCCCGTGCTCGTCATCTCTGTCTGCTACAGCCTCATGATCCGGCGGCTCCGTGGAGTCCG  914

Qy      685  CCAGCAGCAAGCAACTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGT  744
             || ||     |   |      |       || | |   | ||| |        |||| ||
Db      915  CCTGCTCTCGGGCTCCCGAGAGAAGGACCGGAACCTGCGGCGCATCACTCGGCTGGTGCT  974

Qy      745  TGTGATCTTCTCTATACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTC  804
             |||| |   |   || |   || |          ||          | |      || |  | | |||
Db      975  GGTGGTAGTGGCTGTGTTCGTGGGCTGCTGGACGCCTGTCCAGGTCTTCGTGCTGGCCCA  1034

Qy      805  ACGCCTGGATAGTTGGCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACT  864
             | | ||||     | |||     |   |    ||    ||||||          | | |||
Db     1035  AGGGCTGGGGGTTCAGCCGAGCAGCGAGACTGCCGTGGCCATTCTGCGCTTCTGCAC---  1091

Qy      865  GACACGGCCTCTGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGG  924
               | | |||| ||   | |||||     | |||| ||||| ||||| | || |||
Db     1092  ------GGCCCTGGGCTACGTCAACAGCTGCCTCAACCCCATCCTCTACGCCTTCCTGGA  1145

Qy      925  AGACCATTACA  935
             ||   |  |  ||
Db     1146  TGAGAACTTCA  1156
```

RESULT 10
US-09-023-655-1417
; Sequence 1417, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
;   APPLICANT: Cocks, Benjamin G.
;   APPLICANT: Susan G. Stuart
;   APPLICANT: Jeffrey J. Seilhamer
;   TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
;   TITLE OF INVENTION: EXPRESSION
;   NUMBER OF SEQUENCES: 1508
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 PORTER DRIVE
;   CITY: PALO ALTO
;   STATE: CALIFORNIA
;   COUNTRY: USA
;   ZIP: 94304
;   COMPUTER READABLE FORM:

```
    Query Match              5.3%;  Score 82.2;  DB 4;  Length 1973;
    Best Local Similarity   44.5%;  Pred. No. 4e-13;
    Matches  379;  Conservative   0;  Mismatches  463;  Indels   9;  Gaps   1;


Qy         85 GGCTATCTTGAATAAGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACT 144
              |  |  || |     ||| |  | ||   |  | ||      || ||  |   || ||
Db        315 GCCCCTCGGGCTCAAGGTCACCATCGTGGGGCTCTACCTGGCCGTGTGTGTCGGAGGGCT 374


Qy        145 GCTTGGGAATGTCACTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAG 204
              || |||||   |  ||| ||| || | ||||      ||  || |     ||     ||
Db        375 CCTGGGGAACTGCCTTGTCATGTACGTCATCCTCAGGCACACCAAAATGAAGACAGCCAC 434


Qy        205 CAATGTCTATCTTTTTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCAT 264
              |||| | || | |||||||| ||| | | |||  | ||||| || || || ||| |
Db        435 CAATATTTACATCTTTAACCTGGCCCTGGCCGACACTCTGGTCCTGCTGACGCTGCCCTT 494


Qy        265 CCTGATAAAGAGTTATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAA 324
              || |  | |        |  |    || | | ||| || || || | | | |
Db        495 CCAGGGCACGGACATCCTCCTGGGCTTCTGGCCGTTTGGGAATGCGCTGTGCAAGACAGT 554


Qy        325 CCGATATGTGCTTCACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCAT 384
              |        |    || |||| | | ||||||||| | ||  ||| ||| ||| || |
Db        555 CATTGCCATTGACTACTACAACATGTTCACCAGCACCTTCACCCTAACTGCCATGAGTGT 614


Qy        385 GGACCGATATCTGCTCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATT 444
              ||| || ||| |   |||    |||| |||| |  | |   | | |   |  |
Db        615 GGATCGCTATGTAGCCATCTGCCACCCCATCCGTGCCCTCGACGTCCGCACGTCCAGCAA 674
```

```
Qy      445  TGCCATTTTAATCTCGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCT  504
             |||        ||     ||||  |||||||| | |     | | |  | | | ||| |
Db      675  AGCCCAGGCTGTCAATGTGGCCATCTGGGCCCTGGCCTCTGTTGTCGGTGTTCCCGTTGC  734

Qy      505  CACTTTCATCAATTCTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTC  564
             ||   |   |    |       | || || |  |  | ||| | |    |        | |
Db      735  CATCATGGGCTCGGCACAGGTCGAGGATGAAGAGATCGAGTGCCTGGTGGAGATCCCTAC  794

Qy      565  TGGAAACCCTGAACACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAAT  624
               |   | |         |   | |   | |||||| |     | ||   |||| |   |
Db      795  CCCTCAGGATTACTGGGGCCCGGTGTTTGCCATCTGCATCTTCCTCTTCTCCTTCATCGT  854

Qy      625  TCCTCTCTGTGATGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAG  684
             ||  |    || || |   ||| ||||  |   ||| |       |     |       |
Db      855  CCCCGTGCTCGTCATCTCTGTCTGCTACAGCCTCATGATCCGGCGGCTCCGTGGAGTCCG  914

Qy      685  CCAGCAGCAAGCAACTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGT  744
             || ||   |   |    |         || |  |   | ||| |        |||| | |
Db      915  CCTGCTCTCGGGCTCCCGAGAGAAGGACCGGAACCTGCGGCGCATCACTCGGCTGGTGCT  974

Qy      745  TGTGATCTTCTCTATACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTC  804
             |||  |  |  || |  || |  ||       ||      | |     || | | |||
Db      975  GGTGGTAGTGGCTGTGTTCGTGGGCTGCTGGACGCCTGTCCAGGTCTTCGTGCTGGCCCA  1034

Qy      805  ACGCCTGGATAGTTGGCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACT  864
             | | ||||     |   |||      |     ||     ||||||     | | |||
Db     1035  AGGGCTGGGGGTTCAGCCGAGCAGCGAGACTGCCGTGGCCATTCTGCGCTTCTGCAC---  1091

Qy      865  GACACGGCCTCTGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGG  924
               | ||||| ||    | |||||  | ||||| |||| |||||| ||||| | || |||
Db     1092  ------GGCCCTGGGCTACGTCAACAGCTGCCTCAACCCCATCCTCTACGCCTTCCTGGA  1145

Qy      925  AGACCATTACA  935
             ||  | | ||
Db     1146  TGAGAACTTCA  1156
```

RESULT 11
US-09-976-594-171
; Sequence 171, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH
STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 171
;   LENGTH: 3205
;   TYPE: DNA

    Query Match   5.3%;   Score 82.2;   DB 4;   Length 3205;
    Best Local Similarity   44.5%;   Pred. No. 5.3e-13;
    Matches 379; Conservative   0; Mismatches 463; Indels   9; Gaps   1;

```
Qy        85 GGCTATCTTGAATAAGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACT 144
             | |  || |   ||| | | || | | || |  || || | | || ||
Db       389 GCCCCTCGGGCTCAAGGTCACCATCGTGGGGCTCTACCTGGCCGTGTGTGTCGGAGGGCT 448

Qy       145 GCTTGGGAATGTCACTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAG 204
             || ||||||   |  ||| ||| || |  ||||     || ||| |   ||   ||
Db       449 CCTGGGGAACTGCCTTGTCATGTACGTCATCCTCAGGCACACCAAAATGAAGACAGCCAC 508

Qy       205 CAATGTCTATCTTTTTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCAT 264
             |||| | || | ||||||||| || | | || |  |||||| || || ||| |
Db       509 CAATATTTACATCTTTAACCTGGCCCTGGCCGACACTCTGGTCCTGCTGACGCTGCCCTT 568

Qy       265 CCTGATAAAGAGTTATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAA 324
             || |   | |        |  |   || | | | | ||| ||| || || | |  |
Db       569 CCAGGGCACGGACATCCTCCTGGGCTTCTGGCCGTTTGGGAATGCGCTGTGCAAGACAGT 628

Qy       325 CCGATATGTGCTTCACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCAT 384
             |      |   || |||| | | |||||||||| | || ||| ||| ||| || |
Db       629 CATTGCCATTGACTACTACAACATGTTCACCAGCACCTTCACCCTAACTGCCATGAGTGT 688

Qy       385 GGACCGATATCTGCTCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATT 444
             ||| || ||| |   |||    |||| |||| |  | | |  |  |
Db       689 GGATCGCTATGTAGCCATCTGCCACCCCATCCGTGCCCTCGACGTCCGCACGTCCAGCAA 748

Qy       445 TGCCATTTTAATCTCGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCT 504
             |||        ||    |||| |||||||||| | |  | | | | | | | ||| |
Db       749 AGCCCAGGCTGTCAATGTGGCCATCTGGGCCCTGGCCTCTGTTGTCGGTGTTCCCGTTGC 808

Qy       505 CACTTTCATCAATTCTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTC 564
             ||  |  |  |      | || || |  | | ||| | |     | |
Db       809 CATCATGGGCTCGGCACAGGTCGAGGATGAAGAGATCGAGTGCCTGGTGGAGATCCCTAC 868

Qy       565 TGGAAACCCTGAACACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAAT 624
              |  | |    |    | |  |  | ||||| | || ||| | |
Db       869 CCCTCAGGATTACTGGGGCCCGGTGTTTGCCATCTGCATCTTCCTCTTCTCCTTCATCGT 928

Qy       625 TCCTCTCTCTGTGATGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAG 684
             || |   || || |  ||| |||| |  | ||| |     |    |       |
Db       929 CCCCGTGCTCGTCATCTCTGTCTGCTACAGCCTCATGATCCGGCGGCTCCGTGGAGTCCG 988

Qy       685 CCAGCAGCAAGCAACTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGT 744
             || ||     |    |        |        ||| | | ||| |     |||| | |
Db       989 CCTGCTCTCGGGCTCCCGAGAGAAGGACCGGAACCTGCGGCGCATCACTCGGCTGGTGCT 1048

Qy       745 TGTGATCTTCTCTATACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTC 804
             ||| |  |  || | || |  ||    ||       | |     || | | |||
Db      1049 GGTGGTAGTGGCTGTGTTCGTGGGCTGCTGGACGCCTGTCCAGGTCTTCGTGCTGGCCCA 1108
```

```
Qy      805 ACGCCTGGATAGTTGGCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACT 864
            | | |||      |   |||     |     ||      ||||||       | | |||
Db     1109 AGGGCTGGGGGGTTCAGCCGAGCAGCGAGACTGCCGTGGCCATTCTGCGCTTCTGCAC--- 1165

Qy      865 GACACGGCCTCTGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGG 924
            | | |||| ||    | |||||    | |||| |||||| |||||  | || |||
Db     1166 ------GGCCCTGGGCTACGTCAACAGCTGCCTCAACCCCATCCTCTACGCCTTCCTGGA 1219

Qy      925 AGACCATTACA 935
            ||  | | ||
Db     1220 TGAGAACTTCA 1230


RESULT 12
US-09-023-655-992
; Sequence 992, Application US/09023655
; Patent No. 6607879
;  GENERAL INFORMATION:
;    APPLICANT:  Cocks, Benjamin G.
;    APPLICANT:  Susan G. Stuart
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
;    TITLE OF INVENTION:  EXPRESSION
;    NUMBER OF SEQUENCES:  1508
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/023,655
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0001 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  992:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1158 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
```

    Query Match   5.2%;   Score 80; DB 4; Length 1158;
    Best Local Similarity   47.3%;   Pred. No. 1.2e-12;
    Matches 276; Conservative   0; Mismatches 305; Indels   3; Gaps   1;

```
Qy         98 AAGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTC 157
              ||||     ||| ||| |||||||    || ||| | || | || || || ||     |
Db        171 AAGTTGCTCCTTGCTGTCTTTTATTGCCTCCTGTTTGTATTCAGTCTTCTGGGAAACAGC 230

Qy        158 ACTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTT 217
               || | | ||   |  |||||| ||||    ||| || ||   |||| || ||
Db        231 CTGGTCATCCTGGTCCTTGTGGTCTGCAAGAAGCTGAGGAGCATCACAGATGTATACCTC 290

Qy        218 TTTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGT 277
              || ||||||  || | ||||||| |   ||| | | || ||| | || | ||      |
Db        291 TTGAACCTGGCCCTGTCTGACCTGCTTTTTGTCTTCTCCTTCCCCTTTCAGACCTA---C 347

Qy        278 TATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTT 337
              |||        || || ||   | |||   ||| | || |                |  |
Db        348 TATCTGCTGGACCAGTGGGTGTTTGGGACTGTAATGTGCAAAGTGGTGTCTGGCTTTTAT 407

Qy        338 CACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTG 397
              |||     | |||||||| ||||||  | ||| |||| |||| || ||||| | || |||
Db        408 TACATTGGCTTCTACAGCAGCATGTTTTTCATCACCCTCATGAGTGTGGACAGGTACCTG 467

Qy        398 CTCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATC 457
               |     | | | |   | |  | |     | |   || ||   | |  || |      |
Db        468 GCTGTTGTCCATGCCGTGTATGCCCTAAAGGTGAGGACGATCAGGATGGGCACAACGCTG 527

Qy        458 TCGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAAT 517
              | |||||| || |||    |  |    | || |   | ||  |||| || |  |  | |
Db        528 TGCCTGGCAGTATGGCTAACCGCCATTATGGCTACCATCCCATTGCTAGTGTTTTACCAA 587

Qy        518 TCTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAA 577
              | |   ||||   |     ||| ||   | ||   | || |   | |     | |
Db        588 GTGGCCTCTGAAGATGGTGTTCTACAGTGTTATTCATTTTACAATCAACAGACTTTGAAG 647

Qy        578 CACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTG 637
              || || | |   || | ||  |||   | | ||||||| | || || || || || | |
Db        648 TGGAAGATCTTCACCAACTTCAAAATGAACATTTTAGGCTTGTTGATCCCATTCACCATC 707

Qy        638 ATGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAG 681
              |     |||| |||| || || || || |     | |||||| |
Db        708 TTTATGTTCTGCTACATTAAAATCCTGCACCAGCTGAAGAGGTG 751
```

    Query Match   5.2%;   Score 80;   DB 1;   Length 1586;
    Best Local Similarity   47.3%;   Pred. No. 1.5e-12;
    Matches 276; Conservative   0; Mismatches 305; Indels   3; Gaps   1;

Qy   98 AAGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTC 157
             |||| ||| ||| |||||| || ||| | || | || || || || |
Db   533 AAGTTGCTCCTTGCTGTCTTTTATTGCCTCCTGTTTGTATTCAGTCTTCTGGGAAACAGC 592

Qy   158 ACTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTT 217
             || | | ||| | |||||| |||| ||| || || || |||| || ||
Db   593 CTGGTCATCCTGGTCCTTGTGGTCTGCAAGAAGCTGAGGAGCATCACAGATGTATACCTC 652

Qy   218 TTTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGT 277
             || |||||| || | ||||||| | ||| | | | || | ||| | || | | || |
Db   653 TTGAACCTGGCCCTGTCTGACCTGCTTTTTGTCTTCTCCTTCCCCTTTCAGACCTA---C 709

```
Qy    278 TATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTT 337
          |||         || || ||   | |||    ||| | || |             |  |
Db    710 TATCTGCTGGACCAGTGGGTGTTTGGGACTGTAATGTGCAAAGTGGTGTCTGGCTTTTAT 769

Qy    338 CACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTG 397
          |||      | |||||| ||||||  | ||| ||||  |||| ||  |||||| | || |||
Db    770 TACATTGGCTTCTACAGCAGCATGTTTTTCATCACCCTCATGAGTGTGGACAGGTACCTG 829

Qy    398 CTCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATC 457
           |    | | |   | |      | |   | ||     | ||    | |  |       |
Db    830 GCTGTTGTCCATGCCGTGTATGCCCTAAAGGTGAGGACGATCAGGATGGGCACAACGCTG 889

Qy    458 TCGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAAT 517
          | |||||| || |||       |   |  | |   | ||| ||||    ||   | |
Db    890 TGCCTGGCAGTATGGCTAACCGCCATTATGGCTACCATCCCATTGCTAGTGTTTTACCAA 949

Qy    518 TCTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAA 577
           | |  ||||| |   |     | ||    | | | |    |       |       |
Db    950 GTGGCCTCTGAAGATGGTGTTCTACAGTGTTATTCATTTTACAATCAACAGACTTTGAAG 1009

Qy    578 CACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTG 637
          ||  || |   || | ||    |||   | || ||||| | || ||  || ||  |   |
Db   1010 TGGAAGATCTTCACCAACTTCAAAATGAACATTTTAGGCTTGTTGATCCCATTCACCATC 1069

Qy    638 ATGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAG 681
          |     |||| ||||   || ||   |     | |||||| |
Db   1070 TTTATGTTCTGCTACATTAAAATCCTGCACCAGCTGAAGAGGTG 1113
```

RESULT 14
US-09-016-434-1096
; Sequence 1096, Application US/09016434
; Patent No. 6500938
;   GENERAL INFORMATION:
;   APPLICANT: Janice Au-Young
;   APPLICANT: Jeffrey J. Seilhamer
;   TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
;   TITLE OF INVENTION: PATHWAY GENE EXPRESSION
;   NUMBER OF SEQUENCES: 1490
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 PORTER DRIVE
;   CITY: PALO ALTO
;   STATE: CALIFORNIA
;   COUNTRY: USA
;   ZIP: 94304
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/016,434
;   FILING DATE: HEREWITH
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:

```
  Query Match               5.2%;   Score 80;   DB 4;   Length 1953;
  Best Local Similarity   47.3%;   Pred. No. 1.7e-12;
  Matches  276;  Conservative    0;  Mismatches  305;  Indels    3;  Gaps    1;


Qy        98 AAGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTC 157
             ||||     |||  |||  ||||||     ||  |||  | ||  | || ||  || ||      |
Db       369 AAGTTGCTCCTTGCTGTCTTTTATTGCCTCCTGTTTGTATTCAGTCTTCTGGGAAACAGC 428


Qy       158 ACTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTT 217
              ||  | | |  |     |  ||||||  ||||     ||| || ||    |||| || ||
Db       429 CTGGTCATCCTGGTCCTTGTGGTCTGCAAGAAGCTGAGGAGCATCACAGATGTATACCTC 488


Qy       218 TTTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGT 277
              || |||||| || | ||||||| |   ||| | | | || | ||| | | ||    |
Db       489 TTGAACCTGGCCCTGTCTGACCTGCTTTTTGTCTTCTCCTTCCCCTTTCAGACCTA---C 545


Qy       278 TATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTT 337
              |||        ||  || ||   | |||    |||  | ||| |                | |
Db       546 TATCTGCTGGACCAGTGGGTGTTTGGGACTGTAATGTGCAAAGTGGTGTCTGGCTTTTAT 605


Qy       338 CACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTG 397
              |||      | |||||||| ||||||  | ||| ||||  |||| || ||||| | || |||
Db       606 TACATTGGCTTCTACAGCAGCATGTTTTTCATCACCCTCATGAGTGTGGACAGGTACCTG 665


Qy       398 CTCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATC 457
               |     |  |  |   | |    |  |    |  | ||    | ||   | | ||       |
Db       666 GCTGTTGTCCATGCCGTGTATGCCCTAAAGGTGAGGACGATCAGGATGGGCACAACGCTG 725


Qy       458 TCGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAAT 517
              |  |||||| || |||    |  |  |    |  | ||   | ||  ||||  ||    | |
Db       726 TGCCTGGCAGTATGGCTAACCGCCATTATGGCTACCATCCCATTGCTAGTGTTTTACCAA 785


Qy       518 TCTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAA 577
              | | | |||| |  |   |    ||| |     | |    | |   | |        |      |
Db       786 GTGGCCTCTGAAGATGGTGTTCTACAGTGTTATTCATTTTACAATCAACAGACTTTGAAG 845


Qy       578 CACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTG 637
```

```
                           ||  || |   || | ||      |||    | || |||||  | || || || |  |
Db      846 TGGAAGATCTTCACCAACTTCAAAATGAACATTTTAGGCTTGTTGATCCCATTCACCATC 905

Qy      638 ATGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAG 681
                |    |||| ||||   || ||   |    |    | |||||| |
Db      906 TTTATGTTCTGCTACATTAAAATCCTGCACCAGCTGAAGAGGTG 949
```

RESULT 15
US-09-023-655-955
; Sequence 955, Application US/09023655
; Patent No. 6607879
;   GENERAL INFORMATION:
;   APPLICANT: Cocks, Benjamin G.
;   APPLICANT: Susan G. Stuart
;   APPLICANT: Jeffrey J. Seilhamer
;   TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
;   TITLE OF INVENTION: EXPRESSION
;   NUMBER OF SEQUENCES: 1508
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 PORTER DRIVE
;   CITY: PALO ALTO
;   STATE: CALIFORNIA
;   COUNTRY: USA
;   ZIP: 94304
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/023,655
;   FILING DATE: HEREWITH
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Zeller, Karen J.
;   REGISTRATION NUMBER: 37,071
;   REFERENCE/DOCKET NUMBER: PA-0001 US
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (650) 855-0555
;   TELEFAX: (650) 845-4166
;   INFORMATION FOR SEQ ID NO: 955:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2608 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   IMMEDIATE SOURCE:
;   LIBRARY: GENBANK
;   CLONE: g1468978
US-09-023-655-955

```
Query Match              5.2%;   Score 80;  DB 4;  Length 2608;
Best Local Similarity   47.3%;   Pred. No. 2e-12;
  Matches  276;  Conservative    0;  Mismatches 305;  Indels    3;  Gaps     1;

Qy       98 AAGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTC 157
            ||||    |||   |||  ||||||    ||   |||  |  ||   |  ||  ||  ||   |
Db      463 AAGTTGCTCCTTGCTGTCTTTTATTGCCTCCTGTTTGTATTCAGTCTTCTGGGAAACAGC 522

Qy      158 ACTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTT 217
              ||  |  |  |||       |   |||||| ||||    ||| ||  ||    ||||  || ||
Db      523 CTGGTCATCCTGGTCCTTGTGGTCTGCAAGAAGCTGAGGAGCATCACAGATGTATACCTC 582

Qy      218 TTTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGT 277
             || |||||| |||  | |||||||  |   |||  |  |  |  ||  |  ||| |  |  ||     |
Db      583 TTGAACCTGGCCCTGTCTGACCTGCTTTTTGTCTTCTCCTTCCCCTTTCAGACCTA---C 639

Qy      278 TATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTT 337
            |||        ||  || || ||   | |||   |||   | || |               |    |
Db      640 TATCTGCTGGACCAGTGGGTGTTTGGGACTGTAATGTGCAAAGTGGTGTCTGGCTTTTAT 699

Qy      338 CACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTG 397
            |||    | ||||||| |||||| | ||| |||| |||| ||  ||||| | || |||
Db      700 TACATTGGCTTCTACAGCAGCATGTTTTTCATCACCCTCATGAGTGTGGACAGGTACCTG 759

Qy      398 CTCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATC 457
             |      |  |  |    |  |      |   |   |    | ||       | | ||        |
Db      760 GCTGTTGTCCATGCCGTGTATGCCCTAAAGGTGAGGACGATCAGGATGGGCACAACGCTG 819

Qy      458 TCGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAAT 517
            |  ||||| ||  |||        |  |   |  |      | ||  ||||    ||   | |
Db      820 TGCCTGGCAGTATGGCTAACCGCCATTATGGCTACCATCCCATTGCTAGTGTTTTACCAA 879

Qy      518 TCTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAA 577
            | | |  |||| |   |  | ||     |  ||   | |  | |        |       |
Db      880 GTGGCCTCTGAAGATGGTGTTCTACAGTGTTATTCATTTACAATCAACAGACTTTGAAG 939

Qy      578 CACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTG 637
            || || | |  || | || |||   | ||||||| | || || || || || |  |
Db      940 TGGAAGATCTTCACCAACTTCAAAATGAACATTTTAGGCTTGTTGATCCCATTCACCATC 999

Qy      638 ATGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAG 681
            |    |||| ||||   || ||  |    | |||||| |
Db     1000 TTTATGTTCTGCTACATTAAAATCCTGCACCAGCTGAAGAGGTG 1043


Search completed: August 24, 2004, 16:05:17
Job time : 128 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 14:51:16 ; Search time 749 Seconds
                                           (without alignments)
                                           10119.388 Million cell updates/sec

Title: US-09-891-138A-1
Perfect score: 1543
Sequence: 1 gctcctggcagagttttctg..........tgcctaaataaatcaatata 1543

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 3228839 seqs, 2456066551 residues

Total number of hits satisfying chosen parameters: 6457678

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Published_Applications_NA:*
                 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
                 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
                 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
                 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
                 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
                 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
                 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
                 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
                 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
                 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
                 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
                 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
                 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
                 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
                 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
                 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
                 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
                 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
                 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

         Pred. No. is the number of results predicted by chance to have a
         score greater than or equal to the score of the result being printed,
         and is derived by analysis of the total score distribution.

                                 SUMMARIES
                      %
Result   Query

No.	Score	Match	Length	DB	ID	Description
1	1543	100.0	1543	10	US-09-891-138A-1	Sequence 1, Appli
2	592.4	38.4	1005	13	US-09-875-076-35	Sequence 35, Appl
3	592.4	38.4	1005	13	US-09-876-252-37	Sequence 37, Appl
4	592.4	38.4	1005	15	US-10-272-983-35	Sequence 35, Appl
5	592.4	38.4	1005	15	US-10-393-807-35	Sequence 35, Appl
6	592.4	38.4	1005	16	US-10-417-820A-37	Sequence 37, Appl
7	592.4	38.4	1005	17	US-10-723-955-37	Sequence 37, Appl
8	592.4	38.4	1005	17	US-10-782-596-35	Sequence 35, Appl
9	592.4	38.4	1380	15	US-10-225-567A-566	Sequence 566, App
10	592.4	38.4	1436	10	US-09-764-886-36	Sequence 36, Appl
11	592.4	38.4	1436	13	US-09-764-886-36	Sequence 36, Appl
12	592.4	38.4	1436	16	US-10-264-237-1352	Sequence 1352, Ap
13	592.4	38.4	1542	12	US-10-311-671-20	Sequence 20, Appl
14	592.4	38.4	4232	10	US-09-764-886-11	Sequence 11, Appl
15	592.4	38.4	4232	13	US-09-764-886-11	Sequence 11, Appl
16	590.8	38.3	1428	15	US-10-270-587-1	Sequence 1, Appli
17	589.2	38.2	1996	17	US-10-706-532-1	Sequence 1, Appli
c 18	158.8	10.3	639	17	US-10-283-975A-793	Sequence 793, App
19	126.6	8.2	1014	9	US-09-943-798-3	Sequence 3, Appli
20	126.6	8.2	1014	10	US-09-885-453-2	Sequence 2, Appli
21	126.6	8.2	1014	13	US-10-344-728-5	Sequence 5, Appli
22	126.6	8.2	1014	13	US-10-343-650A-13	Sequence 13, Appl
23	126.6	8.2	1014	15	US-10-270-144-1	Sequence 1, Appli
24	126.6	8.2	1014	15	US-10-188-405-7	Sequence 7, Appli
25	126.6	8.2	1014	15	US-10-079-384-13	Sequence 13, Appl
26	126.6	8.2	1014	15	US-10-225-567A-646	Sequence 646, App
27	126.6	8.2	1014	15	US-10-321-807-27	Sequence 27, Appl
28	126.6	8.2	1014	17	US-10-321-807-27	Sequence 27, Appl
29	126.6	8.2	1014	17	US-10-314-048A-27	Sequence 27, Appl
30	126.6	8.2	1081	13	US-10-375-157-1	Sequence 1, Appli
31	126.6	8.2	1081	15	US-10-010-568-1	Sequence 1, Appli
32	126.6	8.2	1414	15	US-10-017-161-525	Sequence 525, App
33	126.6	8.2	1414	16	US-10-292-798-463	Sequence 463, App
34	126.6	8.2	1729	13	US-10-400-991-17	Sequence 17, Appl
35	126.6	8.2	9905	15	US-10-270-144-3	Sequence 3, Appli
36	125	8.1	1014	15	US-10-023-775B-1	Sequence 1, Appli
37	125	8.1	1014	15	US-10-278-141-10	Sequence 10, Appl
38	125	8.1	1014	16	US-10-296-081-10	Sequence 10, Appl
39	125	8.1	1014	17	US-10-763-854-1	Sequence 1, Appli
40	125	8.1	1560	16	US-10-023-634-5	Sequence 5, Appli
41	125	8.1	1851	17	US-10-055-569A-7	Sequence 7, Appli
42	119.2	7.7	1020	10	US-09-782-974C-85	Sequence 85, Appl
43	119	7.7	1313	9	US-09-728-422-1	Sequence 1, Appli
44	104	6.7	1020	15	US-10-023-586B-3	Sequence 3, Appli
45	104	6.7	1020	17	US-10-763-972-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-891-138A-1
; Sequence 1, Application US/09891138A
; Publication No. US20030083245A1
; GENERAL INFORMATION:

```
  Query Match              100.0%;  Score 1543;  DB 10;  Length 1543;
  Best Local Similarity    100.0%;  Pred. No. 0;
  Matches 1543;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 GCTCCTGGCAGAGTTTTCTGTCGAGACAGAAGCCGACAGCAGAATGGCACAGAATTTATC 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 GCTCCTGGCAGAGTTTTCTGTCGAGACAGAAGCCGACAGCAGAATGGCACAGAATTTATC 60

Qy        61 TTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTA 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 TTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATAAGTACTACCTCTCTGCATTTTA 120

Qy       121 TGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGTGTTCGGCTACCTCTT 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 TGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCACTGTGGTGTTCGGCTACCTCTT 180

Qy       181 CTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCTTTCCATCTCTGACTT 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 CTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTTTTAACCTTTCCATCTCTGACTT 240

Qy       241 TGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAATGATAAGGGGACCTA 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 TGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTTATGCCAATGATAAGGGGACCTA 300

Qy       301 TGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAACCTCTACACCAGCAT 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 TGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTCACACCAACCTCTACACCAGCAT 360

Qy       361 CCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGCTCATGAAGTACCCTTTCCGAGA 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 CCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGCTCATGAAGTACCCTTTCCGAGA 420

Qy       421 ACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCTCGCTGGCTGTCTGGGCCTTAGT 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    421  ACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCTCGCTGGCTGTCTGGGCCTTAGT  480

Qy    481  GACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATTCTGTCCCAAAAGAAGAGGGCAG  540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  GACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATTCTGTCCCAAAAGAAGAGGGCAG  540

Qy    541  TAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAACACAATCTCATTTACAGCCTCTG  600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  TAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAACACAATCTCATTTACAGCCTCTG  600

Qy    601  CCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGATGTGCTTCTTCTACTACAAGAT  660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  CCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGATGTGCTTCTTCTACTACAAGAT  660

Qy    661  GGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAACTGCCCTGCCACTGGACAAACC  720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  GGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAACTGCCCTGCCACTGGACAAACC  720

Qy    721  CCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTATACTCTTCACACCCTATCATAT  780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  CCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTATACTCTTCACACCCTATCATAT  780

Qy    781  CATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTTGGCCACAAGGATGTACACAGAA  840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  CATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTTGGCCACAAGGATGTACACAGAA  840

Qy    841  GGCCATCAAATCTATATACACACTGACACGGCCTCTGGCCTTTCTGAACAGTGCCATCAA  900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  GGCCATCAAATCTATATACACACTGACACGGCCTCTGGCCTTTCTGAACAGTGCCATCAA  900

Qy    901  TCCCATCTTCTACTTCCTCATGGGAGACCATTACAGAGAGATGCTGATTAGTAAGTTCAG  960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  TCCCATCTTCTACTTCCTCATGGGAGACCATTACAGAGAGATGCTGATTAGTAAGTTCAG  960

Qy    961  ACAATACTTCAAGTCCCTTACATCCTTCAGGACATGAGCTGCTGGATGCAGGTCTTCACT  1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  ACAATACTTCAAGTCCCTTACATCCTTCAGGACATGAGCTGCTGGATGCAGGTCTTCACT  1020

Qy   1021  CAGCCAAAATGAGACACTTGATAAACAGTGCTGTGCAGTTGAGTTTTAACTAAGTAAACC  1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  CAGCCAAAATGAGACACTTGATAAACAGTGCTGTGCAGTTGAGTTTTAACTAAGTAAACC  1080

Qy   1081  ACCATTTCTAGGCTTTAGCTTTCCACCATCCTCCAACCCCCAGGGCTGGAGTACAAGCTG  1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  ACCATTTCTAGGCTTTAGCTTTCCACCATCCTCCAACCCCCAGGGCTGGAGTACAAGCTG  1140

Qy   1141  GGTCCACATGAATCAGAAGGCAGCTCTCTGTTCTGATTTTAGGTTATACCCAGAGTATGG  1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  GGTCCACATGAATCAGAAGGCAGCTCTCTGTTCTGATTTTAGGTTATACCCAGAGTATGG  1200

Qy   1201  AAAAAATAAGGCATGAGAAAGCATTGACATCTTCACTTAAGAACTGAACAAAAGAGAACA  1260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201  AAAAAATAAGGCATGAGAAAGCATTGACATCTTCACTTAAGAACTGAACAAAAGAGAACA  1260

Qy   1261  AATATTGTCAATGTTTGGACACTTAGGATCTGAAATCTTGGAAATTTTAAGACCTCTTTT  1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261  AATATTGTCAATGTTTGGACACTTAGGATCTGAAATCTTGGAAATTTTAAGACCTCTTTT  1320
```

```
Qy    1321 TCTATCAGTGTAAAAGGAATACAAGATAGCTAGTTGCAAATGCTGAATGCATTTCATCAT 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 TCTATCAGTGTAAAAGGAATACAAGATAGCTAGTTGCAAATGCTGAATGCATTTCATCAT 1380

Qy    1381 TGGTCAGGTCGATAAGCGTGTTTCTGAAATAGTCTTATTTTTATTCTTGTAATATTAAAA 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TGGTCAGGTCGATAAGCGTGTTTCTGAAATAGTCTTATTTTTATTCTTGTAATATTAAAA 1440

Qy    1441 TTTATGTGAAAAATGAATATAATTCAATGTACAACATTAGATTTTCTATTTGAAAATTAT 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 TTTATGTGAAAAATGAATATAATTCAATGTACAACATTAGATTTTCTATTTGAAAATTAT 1500

Qy    1501 ATTTCTTGAAAAAATAACTGCTGTGCCTAAATAAATCAATATA 1543
           |||||||||||||||||||||||||||||||||||||||||||
Db    1501 ATTTCTTGAAAAAATAACTGCTGTGCCTAAATAAATCAATATA 1543
```

RESULT 2
US-09-875-076-35
; Sequence 35, Application US/09875076
; Publication No. US20030017528A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Dang, Huong T.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lin, I-Lin
; TITLE OF INVENTION: Human Orphan G Protein Coupled Receptors
; FILE REFERENCE: AREN0050
; CURRENT APPLICATION NUMBER: US/09/875,076
; CURRENT FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 09/417,044
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 60/120,416
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 60/121,851
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/123,946
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,949
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/136,436
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/136,437
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/136,439
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/136,567
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/137,127
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/137,131
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/141,448
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 60/156,653
; PRIOR FILING DATE: 1999-09-29

  Query Match   38.4%;   Score 592.4; DB 13;   Length 1005;
  Best Local Similarity   75.5%;   Pred. No. 2.7e-139;
  Matches 750; Conservative   0; Mismatches 241; Indels   3; Gaps   1;

Qy   39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
             | | ||||||| |||| | |||| | || ||| |||||| ||||||||| | || | |
Db   8 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 67

Qy   99 AGTACTACCTCTCTGCATTTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
             |||||||||| || ||||||| || ||||| || | ||| | ||||| ||| ||
Db   68 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 127

Qy   159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
             ||| || | ||||||| |||||| ||||||||||||||||||||||||| ||| | |||||| |
Db   128 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 187

Qy   219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
             ||||||| || ||||||||||| ||||| |||||||||||| ||||| ||||||| |||||
Db   188 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 247

Qy   279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
             ||||||||| || ||| ||||||| || ||||| |||||||||||||||||||||||||||
Db   248 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 307

Qy   339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
             | ||||||||| |||||||| ||||| ||||||||| || |||||| || |||||| ||
Db   308 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 367

Qy   399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
             | || ||||| |||||||||||||||| |||| ||||||||| || ||||| ||||||||||
Db   368 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCT 427

Qy   459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518

```
              |   ||||  |  ||||   |||||| |||||||||  |  ||||||||| ||   |    |  || |||
Db     428 CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 487

Qy     519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
              ||||    ||  ||  |  ||||  |  ||||  | || | ||||||||||||||| ||||  |
Db     488 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 547

Qy     579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
              ||||  |||||||||||||| |  || || ||  ||||||| ||||||  |||||||| |  |||||
Db     548 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 607

Qy     639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
              ||||  |||||  ||  ||||||||  |    ||||| ||||| ||||| |||||     |||| || |
Db     608 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 667

Qy     699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
              ||||  |||||  ||  ||  ||  || |     |  ||||  ||  ||||  ||  ||  |||||||||||
Db     668 CTGCTCTGCCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 727

Qy     759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
              |  ||  ||  |||||||||||| |||||||| |||| ||||||||||| |||||||||||  ||||
Db     728 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 787

Qy     819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
              |   ||   |    ||  || ||| |  | ||||||  ||  | ||||    ||||||||||||
Db     788 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT 847

Qy     876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
              |||||||||||||||||||||||| |||||||  || |  |||||||| || || |||||||| || | ||
Db     848 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 907

Qy     936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
              | || ||||||||| | | |  | ||||| |||||||| |||||||||||||||| || | ||
Db     908 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 967

Qy     996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA 1029
              | |||  || |  |     |||| |||| ||||
Db     968 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAA 1001
```

RESULT 3
US-09-876-252-37
; Sequence 37, Application US/09876252
; Publication No. US20030018182A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Lehmann-Bruinsma, Karin
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Lowitz, Kevin P.
; APPLICANT: Lin, I-Lin
; APPLICANT: Dang, Huong T.
; APPLICANT: Chen, Ruoping
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: Non-Endogenous Constitively Activated Human G Protein
Coupled Receptors
; FILE REFERENCE: AREN-0054
; CURRENT APPLICATION NUMBER: US/09/876,252

```
;  CURRENT FILING DATE:  2001-06-07
;  PRIOR APPLICATION NUMBER: 09/416,760
;  PRIOR FILING DATE: 1999-10-12
;  PRIOR APPLICATION NUMBER: 09/170,496
;  PRIOR FILING DATE: 1998-10-13
;  PRIOR APPLICATION NUMBER: 60/110,060
;  PRIOR FILING DATE: 1998-11-27
;  PRIOR APPLICATION NUMBER: 60/120,416
;  PRIOR FILING DATE: 1999-02-16
;  PRIOR APPLICATION NUMBER: 60/121,852
;  PRIOR FILING DATE: 1999-02-26
;  PRIOR APPLICATION NUMBER: 60/109,213
;  PRIOR FILING DATE: 1998-11-20
;  PRIOR APPLICATION NUMBER: 60/123,944
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: 60/123,945
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: 60/123,948
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: 60/123,951
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: 60/123,946
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: 60/123,949
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: 60/152,524
;  PRIOR FILING DATE: 1999-09-03
;  PRIOR APPLICATION NUMBER: 60/151,114
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: 60/108,029
;  PRIOR FILING DATE: 1998-11-12
;  PRIOR APPLICATION NUMBER: 60/136,436
;  PRIOR FILING DATE: 1999-05-28
;  PRIOR APPLICATION NUMBER: 60/136,439
;  PRIOR FILING DATE: 1999-05-28
;  PRIOR APPLICATION NUMBER: 60/136,567
;  PRIOR FILING DATE: 1999-05-28
;  PRIOR APPLICATION NUMBER: 60/137,127
;  PRIOR FILING DATE: 1999-05-28
;  PRIOR APPLICATION NUMBER: 60/137,131
;  PRIOR FILING DATE: 1999-05-28
;  PRIOR APPLICATION NUMBER: 60/141,448
;  PRIOR FILING DATE: 1999-06-29
;  PRIOR APPLICATION NUMBER: 60/136,437
;  PRIOR FILING DATE: 1999-05-28
;  PRIOR APPLICATION NUMBER: 60/156,555
;  PRIOR FILING DATE: 1999-09-29
;  PRIOR APPLICATION NUMBER: 60/156,634
;  PRIOR FILING DATE: 1999-09-29
;  PRIOR APPLICATION NUMBER: 60/156,653
;  PRIOR FILING DATE: 1999-09-29
;  PRIOR APPLICATION NUMBER: 60/157,280
;  PRIOR FILING DATE: 1999-10-01
;  PRIOR APPLICATION NUMBER: 60/157,294
;  PRIOR FILING DATE: 1999-10-01
;  PRIOR APPLICATION NUMBER: 60/157,281
;  PRIOR FILING DATE: 1999-10-01
```

```
   Query Match              38.4%;  Score 592.4;  DB 13;  Length 1005;
   Best Local Similarity    75.5%;  Pred. No. 2.7e-139;
   Matches  750;  Conservative    0;  Mismatches  241;  Indels    3;  Gaps    1;

Qy       39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
            | |  ||||||  ||||  | ||||  | || |||  |||||  ||||||||  | || | |
Db        8 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 67

Qy       99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
            |||||||||||| ||      ||||||||  || |||||| || | ||| | |||||| ||| ||
Db       68 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 127

Qy      159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
            ||| || | |||||||| ||||||      |||||||||||||||||||| ||| | |||||| |
Db      128 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 187

Qy      219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
            ||||||| ||  |||||||||| ||||| |||||||||||| |||||| ||||||| |||||
Db      188 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 247

Qy      279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
            |||||||||  ||   |||  ||||||||| || ||||| |||||||||||||||||||||||
Db      248 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 307

Qy      339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
            |  |||||||||| |||||||| ||||||| ||||| ||||||||| || |||||| || |||||  ||
Db      308 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 367

Qy      399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
            | ||  |||||| ||||||||||||||| ||| ||||||||||| || ||||| ||||||||||||
Db      368 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCT 427

Qy      459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
            |  ||||  | |||||  ||||| |||||||||  | |||||||| ||   |  | || |||
Db      428 CCTTGGCCATTTGGGTTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 487

Qy      519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
            ||||    ||  ||  | |||| | ||| |  || | |||||||||||||||| ||||  |
Db      488 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 547

Qy      579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
            |||| ||||||||||||||| | || || || |||||||| ||||| ||||||||| | |||||
Db      548 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 607

Qy      639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
```

```
             |||| ||||| || |||||||| |     ||||| ||||| |||||     |||| || |
Db     608   TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA   667

Qy     699   CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA   758
             |||| ||||| || || || || |      | |||| || |||| || || |||||||||
Db     668   CTGCTCTGCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG   727

Qy     759   TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT   818
             | || || |||||||||||| ||||||| ||| ||||||||||| ||||||||||| ||||
Db     728   TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT   787

Qy     819   G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC   875
             |   || |   || || ||| | ||||||| || | |||| |||||||||||
Db     788   GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT   847

Qy     876   TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA   935
             ||||||||||||||||| ||||||| || ||||||| || || ||||||| || | ||
Db     848   TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA   907

Qy     936   GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT   995
             | || ||||||||| | | | | ||||| ||||||| ||||||||||||||| || | ||
Db     908   GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT   967

Qy     996   GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA   1029
             | ||| || | |     |||| |||| ||||
Db     968   GGGCTCATGAACTCCTACTTTCATTCAGAGAAAA   1001
```

RESULT 4
US-10-272-983-35
; Sequence 35, Application US/10272983
; Publication No. US20030148450A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Dang, Huong T.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lin, I-Lin
; TITLE OF INVENTION: Human Orphan G Protein Coupled Receptors
; FILE REFERENCE: AREN0050
; CURRENT APPLICATION NUMBER: US/10/272,983
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: US/09/417,044
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 60/109,213
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/120,416
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 60/121,851
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/123,946
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,949
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/136,436
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/136,437
; PRIOR FILING DATE: 1999-05-28

```
   Query Match              38.4%;  Score 592.4;  DB 15;  Length 1005;
   Best Local Similarity    75.5%;  Pred. No. 2.7e-139;
   Matches  750;  Conservative    0;  Mismatches  241;  Indels    3;  Gaps    1;

Qy       39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
            | |  ||||||  ||||  | ||||  | || ||| ||||| ||||||||| | || | |
Db        8 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 67

Qy       99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
            |||||||||||| ||    ||||||| || ||||||| || | ||| | ||||| ||| ||
Db       68 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 127

Qy      159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
            ||| || | ||||||| |||||| ||||||     |||||||||||||||||| ||| | ||||| |
Db      128 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 187

Qy      219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
            ||||||| || ||||||||||| ||||| |||||||||||| ||||| ||||||| |||||
Db      188 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 247

Qy      279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
            ||||||||| || |||  ||||||||| || ||||| |||||||||||||||||||||||||
Db      248 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 307

Qy      339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
            |  |||||||||| |||||||| |||||| ||||| ||||||||| || ||||| || ||||| ||
Db      308 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 367

Qy      399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
            | || |||||| |||||||||||||||||||||| |||| |||||||| || ||||| |||||||||||
Db      368 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCT 427

Qy      459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
            |  ||||  | ||||  ||||| ||||||||| | |||||||||| ||  |  | || |||
Db      428 CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 487

Qy      519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
            ||||     ||  ||  | ||||  | ||| | || | || | ||||||||||||||| |||| |
Db      488 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 547

Qy      579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
            |||| |||||||||||| | || || || || ||||||| |||||| ||||||| | |||||
Db      548 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 607
```

```
Qy       639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
             |||| |||||| || |||||||||| |    ||||| ||||| |||||    |||| || |
Db       608 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 667

Qy       699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
             |||| |||||| || || || || |      | |||| || |||| || || ||||||||
Db       668 CTGCTCTGCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 727

Qy       759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
             | || || |||||||||| |||||| ||| |||||||||| ||||||||| |||| ||||
Db       728 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 787

Qy       819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
             |   || |  ||  || ||| |  ||||||| ||| | |||| |||||||||
Db       788 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT 847

Qy       876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
             |||||||||||||||||||| |||||| || |||||||| || || ||||||| || | ||
Db       848 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 907

Qy       936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
             | || |||||||| | | |  | ||||| ||||||| |||||||||||||||| || | ||
Db       908 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 967

Qy       996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA 1029
             | ||| || | |    |||| |||| ||||
Db       968 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAA 1001


RESULT 5
US-10-393-807-35
; Sequence 35, Application US/10393807
; Publication No. US20030175891A1
; GENERAL INFORMATION:
;  APPLICANT: Chen, Ruoping
;  APPLICANT:  Dang, Huong T.
;  APPLICANT:  Liaw, Chen W.
;  APPLICANT:  Lin, I-Lin
;  TITLE OF INVENTION: Human Orphan G Protein Coupled Receptors
;  FILE REFERENCE: AREN0050
;  CURRENT APPLICATION NUMBER: US/10/393,807
;  CURRENT FILING DATE:  2003-03-21
;  PRIOR APPLICATION NUMBER: US/09/417,044
;  PRIOR FILING DATE: 1999-10-12
;  PRIOR APPLICATION NUMBER: 60/109,213
;  PRIOR FILING DATE: 1998-11-20
;  PRIOR APPLICATION NUMBER: 60/120,416
;  PRIOR FILING DATE: 1999-02-16
;  PRIOR APPLICATION NUMBER: 60/121,851
;  PRIOR FILING DATE: 1999-02-26
;  PRIOR APPLICATION NUMBER: 60/123,946
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: 60/123,949
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: 60/136,436
;  PRIOR FILING DATE: 1999-05-28
;  PRIOR APPLICATION NUMBER: 60/136,437
```

```
   Query Match              38.4%;   Score 592.4;   DB 15;   Length 1005;
   Best Local Similarity    75.5%;   Pred. No. 2.7e-139;
   Matches  750;  Conservative    0;  Mismatches  241;  Indels    3;  Gaps    1;

Qy       39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
            | |  ||||||| |||| | |||| | || ||| ||||| |||||||| | || | |
Db        8 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 67

Qy       99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
            |||||||||| || |||||||| || |||||| || | ||| | ||||| ||| ||
Db       68 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 127

Qy      159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
            ||| || | ||||||| |||||| |||||||||||||||||||||| ||| | |||||| |
Db      128 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 187

Qy      219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
            ||||||| || ||||||||||| ||||| |||||||||||| ||||| ||||||| |||||
Db      188 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 247

Qy      279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
            |||||||||| || ||| ||||||||| || ||||| |||||||||||||||||||||||
Db      248 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 307

Qy      339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
            |  |||||||||| ||||||||| |||||| ||||||| || ||||| || ||||| ||
Db      308 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 367

Qy      399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
            | || |||||| ||||||||||||||||||||| |||| ||||||||| || |||||| ||||||||||
Db      368 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCT 427

Qy      459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
            |  |||| | |||| ||||| ||||||||| | ||||||||| || |  | || |||
Db      428 CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 487

Qy      519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
            ||||    ||  ||  | |||| | ||| |  || | |||||||||||||||| ||||  |
Db      488 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 547

Qy      579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
            |||| |||||||||||| | || || || || |||||||| ||||| |||||||| | |||||
Db      548 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 607
```

```
Qy    639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
          |||| |||||| || |||||||| |    ||||| |||||  |||||    ||||  || |
Db    608 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 667

Qy    699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
          |||| |||||| || || || || |    | |||| || |||| || || |||||||||
Db    668 CTGCTCTGCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 727

Qy    759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
          | || || |||||||||||| ||||||| ||| ||||||||||| |||||||||  ||||
Db    728 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 787

Qy    819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
          |   ||  |   ||  || ||| | |||||| || | |||| ||||||||||||
Db    788 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT 847

Qy    876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
          ||||||||||||||||||| |||||| ||  ||||||| || || ||||||| || | ||
Db    848 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 907

Qy    936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
          | || |||||||| | | |  | ||||| ||||||| ||||||||||||||||| || | ||
Db    908 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 967

Qy    996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA 1029
          | ||| || |  |    |||| ||||  ||||
Db    968 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAA 1001


RESULT 6
US-10-417-820A-37
; Sequence 37, Application US/10417820A
; Publication No. US20030229216A1
; GENERAL INFORMATION:
;  APPLICANT: Chen, Ruoping
;  APPLICANT:  Liaw, Chen W.
;  APPLICANT:  Lowitz, Kevin
;  APPLICANT:  Chalmers, Derek T.
;  APPLICANT:  Behan, Dominic P.
;  TITLE OF INVENTION: Constitutively Activated Human G Protein Coupled
;  TITLE OF INVENTION:  Receptors
;  FILE REFERENCE: 7.US28.CON
;  CURRENT APPLICATION NUMBER: US/10/417,820A
;  CURRENT FILING DATE:  2003-04-16
;  PRIOR APPLICATION NUMBER: 09/416,760
;  PRIOR FILING DATE: 1999-10-12
;  PRIOR APPLICATION NUMBER: 09/170,496
;  PRIOR FILING DATE: 1998-10-13
;  PRIOR APPLICATION NUMBER: 60/110,060
;  PRIOR FILING DATE: 1998-11-27
;  PRIOR APPLICATION NUMBER: 60/120,416
;  PRIOR FILING DATE: 1999-02-16
;  PRIOR APPLICATION NUMBER: 60/121,852
;  PRIOR FILING DATE: 1999-02-26
;  PRIOR APPLICATION NUMBER: 60/109,213
;  PRIOR FILING DATE: 1998-11-20
```

```
   Query Match              38.4%;  Score 592.4;  DB 16;  Length 1005;
   Best Local Similarity    75.5%;  Pred. No. 2.7e-139;
   Matches   750;  Conservative    0;  Mismatches  241;  Indels   3;  Gaps    1;

Qy         39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
              | |   ||||||  ||||  | ||||  | || ||| ||||| |||||||| | || | |
Db          8 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 67

Qy         99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
              |||||||||||| ||      |||||||| || |||||| || | ||| | ||||| ||| ||
Db         68 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 127

Qy        159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
              ||| || | ||||||| |||||      |||||||||||||||||||||| ||| | |||||| |
Db        128 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 187

Qy        219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
              ||||||| || ||||||||||| ||||| ||||||||||||| ||||| ||||||| |||||
Db        188 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 247

Qy        279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
              ||||||||| || ||| ||||||||| || ||||| |||||||||||||||||||||||||||||
Db        248 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 307

Qy        339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
              | |||||||||| |||||||| |||||| ||||| ||||||| || ||||| || |||||  ||
Db        308 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 367

Qy        399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
              | || |||||| ||||||||||||||||| |||| ||||||||| || ||||| ||||||||||||
Db        368 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAGAAAGAGTTTGCTATTTTAATCT 427

Qy        459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
              |  |||| | |||| |||||| |||||||| | ||||||||| ||  |  | || |||
Db        428 CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 487

Qy        519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
              ||||     || ||  ||  | ||||  | ||| |  ||  |||||||||||||| |||| |
Db        488 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 547
```

```
Qy     579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
           |||| ||||||||||||| | || || ||  |||||||| ||||| ||||||||| | |||||
Db     548 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 607

Qy     639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
           |||| |||||| || ||||||||| |     ||||| ||||| ||||| |||| || |
Db     608 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 667

Qy     699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
           |||| ||||| || || || || |     | |||| || |||| || || |||||||||
Db     668 CTGCTCTGCCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 727

Qy     759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
           | || || ||||||||||||  ||||||| ||| |||||||||||| |||||||||| ||||
Db     728 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 787

Qy     819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
           |   || |   || || |||  | |||||| || | ||||    |||||||||||
Db     788 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT 847

Qy     876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
           ||||||||||||||||||| |||||| || ||||||| || || ||||||| || | ||
Db     848 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 907

Qy     936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
           | || ||||||||| | | | | ||||| ||||||| ||||||||||||||| || | ||
Db     908 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 967

Qy     996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA 1029
           | ||| || | |    |||| |||| ||||
Db     968 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAA 1001
```

RESULT 7
US-10-723-955-37
; Sequence 37, Application US/10723955
; Publication No. US20040110238A1
; GENERAL INFORMATION:
;   APPLICANT: Behan, Dominic P.
;   APPLICANT: Chalmers, Derek T.
;   APPLICANT: Lin, I-Lin
;   APPLICANT: Liaw, Chen W.
;   APPLICANT: Lehman-Bruinsma, Karin
;   APPLICANT: Lowitz, Kevin P.
;   APPLICANT: Dang, Huong T.
;   APPLICANT: Chen, Ruoping
;   APPLICANT: Gore, Martin
;   APPLICANT: White, Carol
;   TITLE OF INVENTION: Constitutively Activated Human G Protein Coupled
;   TITLE OF INVENTION: Receptors
;   FILE REFERENCE: 7.US29.CON
;   CURRENT APPLICATION NUMBER: US/10/723,955
;   CURRENT FILING DATE: 2003-11-26
;   PRIOR APPLICATION NUMBER: 10/417,820
;   PRIOR FILING DATE: 2003-4-16
;   PRIOR APPLICATION NUMBER: 09/416,760
;   PRIOR FILING DATE: 1999-10-12

```
  Query Match               38.4%;  Score 592.4;  DB 17;  Length 1005;
  Best Local Similarity    75.5%;  Pred. No. 2.7e-139;
  Matches  750;  Conservative   0;  Mismatches  241;  Indels    3;  Gaps     1;


Qy        39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
             | |   ||||||| ||||  | |||| | || ||| ||||| |||||||| | || | |
Db         8 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 67


Qy        99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
             ||||||||||| ||     ||||||| || ||||||| || | ||| | ||||| ||| ||
Db        68 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 127


Qy       159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
             ||| || |  ||||||| |||||| |||||     ||||||||||||||||| ||| | ||||| |
Db       128 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 187


Qy       219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
             ||||||| || |||||||||| ||||| |||||||||| ||||| ||||||| |||||
Db       188 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 247


Qy       279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
             |||||||| || ||| ||||||||| || ||||| |||||||||||||||||||||||||
Db       248 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 307


Qy       339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
             |   |||||||||| |||||||| ||||| ||||||||| || |||||| || |||||| ||
Db       308 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 367


Qy       399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
             | || |||||| |||||||||||||||| |||| |||||||| || ||||| ||||||||||
Db       368 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCT 427
```

```
Qy     459  CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT  518
            |  ||||  |  ||||  |||||  |||||||||  |  |||||||||  ||   |  ||  |||
Db     428  CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC  487

Qy     519  CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC  578
            ||||   ||  ||   |  ||||   |  |||  |   ||  |  ||||||||||||||  ||||   |
Db     488  CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT  547

Qy     579  ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTGTGA  638
            ||||  |||||||||||||  |  ||  ||  ||  |||||||  |||||  |||||||||  |  |||||
Db     548  ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA  607

Qy     639  TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA  698
            ||||  ||||||  ||  |||||||||  |    |||||  |||||    |||||  ||  |
Db     608  TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA  667

Qy     699  CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA  758
            ||||  ||||||  ||  ||  ||  ||  |       |  ||||   ||  ||||  ||  ||  |||||||||
Db     668  CTGCTCTGCCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG  727

Qy     759  TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT  818
            |  ||  ||  ||||||||||||  |||||||  |||  ||||||||||  |||||||||||   ||||
Db     728  TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT  787

Qy     819  G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC  875
            |  ||  |    ||  ||  |||  |    ||||||  ||  |  ||||    |||||||||||
Db     788  GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT  847

Qy     876  TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA  935
            ||||||||||||||||||||  ||||||  ||  |||||||  ||  ||   |||||||  ||  |  ||
Db     848  TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA  907

Qy     936  GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT  995
            |  ||  |||||||||  |  |  |  |  ||||||   |||||||  ||||||||||||||||||  ||  |  ||
Db     908  GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT  967

Qy     996  GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA  1029
            |  |||  ||  |  |        ||||  ||||   ||||
Db     968  GGGCTCATGAACTCCTACTTTCATTCAGAGAAAA  1001
```

RESULT 8
US-10-782-596-35
; Sequence 35, Application US/10782596
; Publication No. US20040137509A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Dang, Huong T.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lin, I-Lin
; TITLE OF INVENTION: Human Orphan G Protein Coupled Receptors
; FILE REFERENCE: AREN0050
; CURRENT APPLICATION NUMBER: US/10/782,596
; CURRENT FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: US/09/875,076
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 09/417,044

```
   Query Match                38.4%;  Score 592.4;  DB 17;  Length 1005;
   Best Local Similarity      75.5%;  Pred. No. 2.7e-139;
   Matches  750;  Conservative    0;  Mismatches  241;  Indels    3;  Gaps    1;

Qy       39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
            | |  ||||||  ||||  |  ||||   |  ||  |||  |||||  |||||||| |  || | |
Db        8 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 67

Qy       99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
            ||||||||||| ||       |||||||  || ||||||  || |  ||| |  ||||| |||  ||
Db       68 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 127

Qy      159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
            ||| || | |||||||| ||||||      |||||||||||||||||||||||| ||| | |||||| |
Db      128 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 187

Qy      219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
            |||||||  ||   |||||||||| ||||| |||||||||| |||||  ||||||||  |||||
Db      188 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 247

Qy      279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
            |||||||||  ||   ||| ||||||||| || |||||| |||||||||||||||||||||||||||
Db      248 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 307

Qy      339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
            |  |||||||||| |||||||| |||||| ||||| |||||||| || ||||| || ||||| ||
Db      308 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 367

Qy      399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
            | ||  |||||  |||||||||||||||| || ||||||| |||||||| || |||||  ||||||||||
Db      368 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCT 427
```

```
Qy    459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
          |  ||||  |  ||||  ||||| ||||||||| | |||||||| || |  | || |||
Db    428 CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 487

Qy    519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
          ||||   ||  || |  |||| | ||| |  || | |||||||||||||| |||| |
Db    488 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 547

Qy    579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
          ||||  |||||||||||  | || || || || ||||||||| ||||| |||||||| |
Db    548 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 607

Qy    639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
          ||||  ||||| || |||||||| |  ||||| ||||| ||||| ||||| |||| || |
Db    608 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 667

Qy    699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
          ||||  ||||| || || || || | |  | ||||  || |||| || ||  |||||||||
Db    668 CTGCTCTGCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 727

Qy    759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
          | || ||  |||||||||||| | |||||||||| |||||||| |||||||||  ||||
Db    728 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 787

Qy    819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
          |  ||  |     || || |||   | |||||| ||  | |||||||||||||||||||
Db    788 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT 847

Qy    876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
          |||||||||||||||||| ||||||  ||  ||||||| || ||  ||||||| || | ||
Db    848 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 907

Qy    936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
          | || |||||||| | | | | ||||| |||||| ||||||||||||||||||| || | ||
Db    908 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 967

Qy    996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA 1029
          | ||| || |  |    |||| |||| ||||
Db    968 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAA 1001
```

RESULT 9
US-10-225-567A-566
; Sequence 566, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED
RECEPTORS (GPCRS)
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144

```
    Query Match              38.4%;  Score 592.4;  DB 15;  Length 1380;
    Best Local Similarity    75.3%;  Pred. No. 3.3e-139;
      Matches  764;  Conservative   0;  Mismatches 246;  Indels   4;  Gaps    2;

Qy        39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
              | |  ||||||  ||||  |  ||||  |  ||  ||| ||||| |||||||| | || | |
Db        50 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 109

Qy        99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
              ||||||||||| ||      |||||||  || ||||||  ||  |  ||| ||||| |||  ||
Db       110 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 169

Qy       159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
              ||| || | ||||||| |||||| |||||      |||||||||||||||||||||| ||| | |||||| |
Db       170 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 229

Qy       219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
              ||||||| ||  ||||||||||| ||||| |||||||||| |||||| ||||||| |||||
Db       230 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 289

Qy       279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
              |||||||||  || ||| ||||||| || ||||| ||||||||||||||||||||||||||
Db       290 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 349

Qy       339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
              |  |||||||||| ||||||||| ||||| ||||||||| || ||||| || |||||  ||
Db       350 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 409

Qy       399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
              | || ||||| |||||||||||||||| ||| ||||||||| || ||||| |||||||||||
Db       410 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCT 469

Qy       459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
              |  ||||  | |||| ||||| ||||||||| || |||||||| ||  |  | || |||
Db       470 CCTTGGCCATTTGGGTTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 529

Qy       519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
              ||||    || || | ||||  | |||  | ||| | || | ||||||||||||| |||| |
Db       530 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 589

Qy       579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
              ||||  |||||||||||| | | ||  | ||  || ||||||| |||||||||  | |||||
Db       590 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 649

Qy       639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
              ||||  ||||  || |||||||| | |||| ||||  |||||| ||||| ||||| || |
Db       650 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 709
```

```
Qy    699  CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
           ||||  ||||||  ||  ||  ||  ||  |      |  ||||  ||  ||||  ||  ||  ||||||||||
Db    710  CTGCTCTGCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 769

Qy    759  TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
           |  ||  ||  |||||||||||    |||||||  |||  ||||||||||  ||||||||||    ||||
Db    770  TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 829

Qy    819  G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
           |    ||    |      ||  ||  |||      |  ||||||  ||  |  ||||      |||||||||||
Db    830  GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT 889

Qy    876  TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
           ||||||||||||||||||||  ||||||  ||  |||||||  ||  ||  |||||||  ||  |  ||
Db    890  TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 949

Qy    936  GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
           |  ||  ||||||||||  |  |  |  |  ||||||    |||||||  |||||||||||||||||  ||  |  ||
Db    950  GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 1009

Qy    996  GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
           |  |||  ||  |  |      ||||  ||||  ||||  ||||      ||||  ||||||
Db   1010  GGGCTCATGAACTCCTACTTTCATTCAGAGAAAGTGAGGGGCTTGTGAAACAG 1063


RESULT 10
US-09-764-886-36
; Sequence 36, Application US/09764886
; Publication No. US20030139327A9
; GENERAL INFORMATION:
;   APPLICANT: Rosen et al.
;   TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;   FILE REFERENCE: PTZ02
;   CURRENT APPLICATION NUMBER: US/09/764,886
;   CURRENT FILING DATE:  2001-01-17
;   Prior application data removed - consult PALM or file wrapper
;   NUMBER OF SEQ ID NOS: 88
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36
;   LENGTH: 1436
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-886-36

   Query Match           38.4%;  Score 592.4;  DB 10;  Length 1436;
   Best Local Similarity   75.3%;  Pred. No. 3.3e-139;
     Matches  764;  Conservative    0;  Mismatches  246;  Indels    4;  Gaps    2;

Qy     39  GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
           |  |    ||||||    ||||    |    ||||    |    ||  |||  ||||||  ||||||||||    |  ||  |    |
Db    100  GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 159

Qy     99  AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
           ||||||||||||  ||        |||||||  ||  ||||||||  ||  |  |||  |  ||||||  |||    ||
Db    160  AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 219

Qy    159  CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
```

```
              |||  || |  |||||||  ||||||    ||||||||||||||||||||| |||  |  ||||||  |
Db     220 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 279

Qy     219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
              ||||||| ||  ||||||||||| |||||| |||||||||||| ||||| ||||||| |||||
Db     280 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 339

Qy     279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
              |||||||||  ||  |||  |||||||| ||  |||||  |||||||||||||||||||||||
Db     340 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 399

Qy     339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
              |  ||||||||| |||||||| |||||| ||||| |||||| || |||||  ||  ||||| ||
Db     400 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 459

Qy     399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
              | ||  |||||  |||||||||||||||| |||| |||||||| || |||||  ||||||||||
Db     460 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCT 519

Qy     459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
              |  ||||  | ||||  ||||| |||||||||  | ||||||||| ||  |  |  || |||
Db     520 CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 579

Qy     519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
              ||||    ||  ||  | |||| | |||  |  || | ||||||||||||||| |||| |
Db     580 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 639

Qy     579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
              ||||  ||||||||||||  |  | ||  || || |||||||| ||||| |||||||| | ||||||
Db     640 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 699

Qy     639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
              ||||  |||||| ||  |||||||| |     ||||| ||||| ||||| |   |||| || |
Db     700 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 759

Qy     699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
              ||||  ||||| ||  || || || | ||  |  | |||| || ||||| || ||  ||||||||||
Db     760 CTGCTCTGCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 819

Qy     759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
              | || ||  |||||||||| |||||||  |||||| ||| |||||||||||| |||||||||   ||||
Db     820 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 879

Qy     819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
              |    ||  |    || || |||   | |||||| ||  | |||||   ||||||||||||||
Db     880 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT 939

Qy     876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
              ||||||||||||||||||||| ||||||| ||  ||||||| || || |||||||| || | ||
Db     940 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 999

Qy     936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
              | || |||||||||| | | | |||||  ||||||| |||||||||||||||||||| || | ||
Db    1000 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 1059

Qy     996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
              | ||| || | | |    |||| |||| |||| ||||    |||| |||||||
```

Db   1060 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAAGTGAGGGGCTTGTGAAACAG 1113

RESULT 11
US-09-764-886-36
; Sequence 36, Application US/09764886
; Publication No. US20020086822A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ02
; CURRENT APPLICATION NUMBER: US/09/764,886
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36
;   LENGTH: 1436
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-886-36

    Query Match   38.4%;   Score 592.4; DB 13;   Length 1436;
    Best Local Similarity   75.3%;   Pred. No. 3.3e-139;
    Matches 764; Conservative   0; Mismatches 246; Indels   4; Gaps   2;

Qy   39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
            | | |||||| |||| | |||| | || ||| |||||| |||||||| | || | |
Db   100 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 159

Qy   99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
            ||||||||||| || ||||||| || |||||| || | ||| | ||||| ||| ||
Db   160 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 219

Qy   159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
            ||| || | ||||||||| ||||| |||||||||||||||||||||| ||| | ||||| |
Db   220 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 279

Qy   219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
            ||||||| || |||||||||| ||||| ||||||||||| ||||| |||||||| |||||
Db   280 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 339

Qy   279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
            ||||||||| || ||| ||||||||| || ||||| ||||||||||||||||||||||||
Db   340 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 399

Qy   339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
            | |||||||||| |||||||| ||||| ||||||||| || |||||| || |||||| ||
Db   400 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 459

Qy   399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
            | || ||||| |||||||||||||||| ||||| ||||||||| || ||||| |||||||||
Db   460 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCT 519

Qy   459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
            | |||| | |||| ||||| ||||||||| | ||||||||| || | | || |||
Db   520 CCTTGGCCATTTGGGTTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 579

```
Qy      519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
                ||||    ||   ||    |  |||| |    | |||  |   ||  |  ||||||||||||||||  ||||    |
Db      580 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 639

Qy      579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
                ||||  ||||||||||||  |  ||  ||  ||    |||||||  |||||  ||||||||||  |  |||||
Db      640 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 699

Qy      639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
                ||||  |||||  ||  ||||||||| |    |||||||  |    ||||||  |||||      ||||    || |
Db      700 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 759

Qy      699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
                ||||  ||||||  ||  ||  ||  ||  ||  |       | ||||   ||  ||||  ||  ||  ||||||||||
Db      760 CTGCTCTGCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 819

Qy      759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
                |  ||  ||  ||||||||||||   |||||||  |||  ||||||||||||  |||||||||||    ||||
Db      820 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 879

Qy      819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
                |    ||   |       ||  ||  |||    |  ||||||||  |  |  ||||        ||||||||||||
Db      880 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT 939

Qy      876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
                ||||||||||||||||||||||||  ||||||  ||    |||||||  ||  ||    ||||||||  ||  |  ||
Db      940 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 999

Qy      936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
                |  ||  ||||||||||  |  |    |  |  |||||   |||||||  |||||||||||||||||||||  |  |  ||
Db     1000 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 1059

Qy      996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
                |  |||  ||  |  |      ||||  ||||    ||||  ||||      |||||  ||||||
Db     1060 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAAGTGAGGGGCTTGTGAAACAG 1113
```

RESULT 12
US-10-264-237-1352
; Sequence 1352, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1352
;   LENGTH: 1436
;   TYPE: DNA

```
     Query Match              38.4%;   Score 592.4;   DB 16;   Length 1436;
     Best Local Similarity    75.3%;   Pred. No. 3.3e-139;
     Matches  764;  Conservative    0;  Mismatches 246;  Indels    4;  Gaps    2;
```

```
Qy        39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
             | |  ||||||  ||||  |  ||||  |  || ||| |||||  ||||||||  | || | |
Db       100 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 159

Qy        99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
             ||||||||||| ||     |||||||  || ||||||| ||  | ||| | ||||| ||| | ||
Db       160 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 219

Qy       159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
             ||| || |  |||||||  ||||||       |||||||||||||||||||||| ||| | |||||| |
Db       220 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 279

Qy       219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
             ||||||| ||  |||||||||||| ||||| |||||||||||| |||||| ||||||| | |||||
Db       280 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 339

Qy       279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
             ||||||||||  || |||  ||||||||| || |||||  ||||||||||||||||||||||||||||
Db       340 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 399

Qy       339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
             |  |||||||||||  |||||||| |||| ||||| |||||||| || ||||| || ||||| | ||
Db       400 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 459

Qy       399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
             | ||  ||||| ||||||||||||||||| |||| |||||||| || |||||  ||||||||||
Db       460 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCT 519

Qy       459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
             |  |||| | |||| |||||| ||||||||||||| | |||||||| || |  | || ||| |
Db       520 CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 579

Qy       519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
             ||||    || || ||| | |||| | ||| | || | |||||||||||||||| |||| |
Db       580 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 639

Qy       579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
             |||| |||||||||||||| | || || || || |||||||| ||||| |||||||| | |||||
Db       640 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 699

Qy       639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
             |||| |||||| || |||||||||| | ||||| ||||| | |||||  |||||    |||| || |
Db       700 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 759

Qy       699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
             |||| ||||| || || || || || | | |||| || |||| || || || ||||||||||
Db       760 CTGCTCTGCCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 819

Qy       759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
             | || || ||||||||||||  ||||||| ||| |||||||||| |||||||||||  ||||
```

```
Db     820 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 879

Qy     819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
               |   || |     || || |||   | |||||| ||  | ||||      |||||||||||
Db     880 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT 939

Qy     876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
               ||||||||||||||||||| |||||| || ||||||||| || || ||||||| || | ||
Db     940 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTCTTTTGGGAGATCACTTCA 999

Qy     936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
               | || ||||||||| | |  |  | |||||   ||||||| |||||||||||||||| || | ||
Db    1000 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 1059

Qy     996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
               | |||  || |  |    |||| ||||  |||| ||||    |||| ||||||
Db    1060 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAGTGAGGGGCTTGTGAAACAG 1113
```

```
  Query Match              38.4%;   Score 592.4;  DB 12;   Length 1542;
  Best Local Similarity    75.3%;   Pred. No. 3.5e-139;
    Matches   764;  Conservative    0;  Mismatches 246;  Indels    4;  Gaps    2;

Qy       39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
            | |   ||||||   ||||   | ||||   | || ||| ||||| |||||||| | || | |
Db      205 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 264

Qy       99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
            |||||||||| ||      ||||||| || ||||||| || | ||| | ||||| ||| ||
Db      265 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 324

Qy      159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
            ||| || | ||||||| |||||| |||||||||||||||||||||| ||| | ||||| |
Db      325 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 384

Qy      219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
            |||||| || ||||||||||| |||||| ||||||||||| ||||| ||||||| |||||
Db      385 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 444

Qy      279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
            |||||||| || ||| |||||||| || ||||| ||||||||||||||||||||||||||
Db      445 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 504

Qy      339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
            |  ||||||||| |||||||||| ||||||| ||||||| || ||||| || ||||| ||
Db      505 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 564

Qy      399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
            | || ||||| |||||||||||||||| |||| |||| ||||||||| || ||||| |||||||||||
Db      565 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCT 624

Qy      459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
```

```
             |   ||||  |  ||||  |||||  ||||||||  |  |||||||| ||  |  |  || |||
Db     625 CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 684

Qy     519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
             ||||    || ||  |  ||||   |  |||  |  ||  | |||||||||||||  ||||  |
Db     685 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 744

Qy     579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA 638
             ||||  ||||||||||||| | ||  ||  ||   |||||||| ||||||  ||||||||| |  |||||
Db     745 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 804

Qy     639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
             ||||  ||||||  ||  |||||||| |     |||||  ||||| |||||     ||||     || |
Db     805 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 864

Qy     699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
             ||||  ||||||  || ||  || ||  |     |  ||||  ||  ||||  || ||  |||||||||||
Db     865 CTGCTCTGCCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 924

Qy     759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
             | || ||  |||||||||  |||||||  |||  ||||||||||| ||||||||||  ||||
Db     925 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 984

Qy     819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
             |    || |      ||  ||  |||   |  ||||||  ||  |  ||||     ||||||||||||
Db     985 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT 1044

Qy     876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
             |||||||||||||||||||||  |||||||  ||||||| || ||   |||||||  || | ||
Db    1045 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 1104

Qy     936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
             |  || |||||||||  |  |  |  ||||||  |||||||  ||||||||||||||  |  ||
Db    1105 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 1164

Qy     996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
             |  |||  ||  |   |      ||||  ||||  ||||  ||||     ||||  ||||||
Db    1165 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAAGTGAGGGGCTTGTGAAACAG 1218
```

RESULT 14
US-09-764-886-11
; Sequence 11, Application US/09764886
; Publication No. US20030139327A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ02
; CURRENT APPLICATION NUMBER: US/09/764,886
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 4232
;   TYPE: DNA
;   ORGANISM: Homo sapiens

Query Match   38.4%; Score 592.4; DB 10; Length 4232;
Best Local Similarity   75.3%; Pred. No. 6.6e-139;
Matches 764; Conservative   0; Mismatches 246; Indels   4; Gaps   2;

```
Qy       39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
            | |  ||||||  ||||  | ||||  | || |||  |||||  |||||||||  | ||  | |
Db      110 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 169

Qy       99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
            ||||||||||| ||    |||||||  || ||||||| ||  | ||| | ||||| ||   ||
Db      170 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 229

Qy      159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
            ||| || |  |||||||| |||||||      ||||||||||||||||||||| ||| | |||||| |
Db      230 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 289

Qy      219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
            ||||||| ||  |||||||||| ||||| ||||||||||| |||||| |||||||| |||||
Db      290 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 349

Qy      279 ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC 338
            |||||||||  ||   ||| |||||||| ||  |||||| |||||||||||||||||||||||
Db      350 ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC 409

Qy      339 ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC 398
            |  |||||||||| ||||||||| |||||| ||||| |||||||| || |||||| || ||||| ||
Db      410 ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA 469

Qy      399 TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT 458
            | || ||||| ||||||||||||||| ||| |||| |||||||| || ||||| |||||||||
Db      470 TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCT 529

Qy      459 CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT 518
            |  |||| | |||| ||||| |||||||| | |||||||| || |  | || |||
Db      530 CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC 589

Qy      519 CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC 578
            ||||    ||  ||  | ||||  | ||| |  || | |||||||||||||| ||||  |
Db      590 CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT 649

Qy      579 ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTGTGA 638
            ||||  |||||||||||||   ||  || || ||  |||||||  |||||||| | |||||
Db      650 ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA 709

Qy      639 TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA 698
            |||| ||||| || |||||||| |  ||||| ||||| |||||  ||||  |||| || |
Db      710 TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA 769

Qy      699 CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA 758
            |||| |||||  || || || || || |  | |||| || ||| || || |||||||||||
Db      770 CTGCTCTGCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG 829

Qy      759 TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT 818
            | || || |||||||||| |  | ||| |||||||||||||||| ||||||||||| ||||
Db      830 TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT 889
```

```
Qy     819 G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC 875
           |   || |    || || |||   | |||||| ||  | ||||   ||||||||||||
Db     890 GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT 949

Qy     876 TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA 935
           ||||||||||||||||||| |||||| || ||||||| || || ||||||| || | ||
Db     950 TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA 1009

Qy     936 GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT 995
           | || ||||||||| | | | ||||| ||||||| |||||||||||||||| || | ||
Db    1010 GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT 1069

Qy     996 GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG 1048
           | |||  || | |    |||| |||| |||| |||| ||||  |||| ||||||
Db    1070 GGGCTCATGAACTCCTACTTTCATTCAGAGAAAAGTGAGGGGCTTGTGAAACAG 1123


RESULT 15
US-09-764-886-11
; Sequence 11, Application US/09764886
; Publication No. US20020086822A1
; GENERAL INFORMATION:
;   APPLICANT: Rosen et al.
;   TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;   FILE REFERENCE: PTZ02
;   CURRENT APPLICATION NUMBER: US/09/764,886
;   CURRENT FILING DATE:  2001-01-17
;   Prior application data removed - consult PALM or file wrapper
;   NUMBER OF SEQ ID NOS: 88
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 4232
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-886-11

   Query Match              38.4%;  Score 592.4;  DB 13;  Length 4232;
   Best Local Similarity    75.3%;  Pred. No. 6.6e-139;
   Matches   764; Conservative   0; Mismatches  246; Indels   4; Gaps   2;

Qy      39 GCAGAATGGCACAGAATTTATCTTGTGAGAATTGGTTGGCAACAGAGGCTATCTTGAATA 98
           | |  ||||||| ||||  | ||||| | || ||| ||||| |||||||||  | ||| |
Db     110 GGATCATGGCATGGAATGCAACTTGCAAAAACTGGCTGGCAGCAGAGGCTGCCCTGGAAA 169

Qy      99 AGTACTACCTCTCTGCATTTTATGCAATCGAGTTCATTTTTGGACTGCTTGGGAATGTCA 158
           |||||||| ||    ||||||| || ||||||| || | ||| | |||| ||| ||| ||
Db     170 AGTACTACCTTTCCATTTTTTATGGGATTGAGTTCGTTGTGGGAGTCCTTGGAAATACCA 229

Qy     159 CTGTGGTGTTCGGCTACCTCTTCTGCATGAAGAACTGGAACAGCAGCAATGTCTATCTTT 218
           ||| || | |||||||| |||||||||||     |||||||||||||||| ||| | |||| |
Db     230 TTGTTGTTTACGGCTACATCTTCTCTCTGAAGAACTGGAACAGCAGTAATATTTATCTCT 289

Qy     219 TTAACCTTTCCATCTCTGACTTTGCTTTCCTGTGCACCCTTCCCATCCTGATAAAGAGTT 278
           ||||||| ||  |||||||||| |||| ||||||||||||  ||||| |||||||| |||||
Db     290 TTAACCTCTCTGTCTCTGACTTAGCTTTTCTGTGCACCCTCCCCATGCTGATAAGGAGTT 349
```

```
Qy       279  ATGCCAATGATAAGGGGACCTATGGAGATGTTCTCTGTATAAGCAACCGATATGTGCTTC  338
              |||||||||  ||   |||   |||||||| ||  |||||| ||||||||||||||||||||
Db       350  ATGCCAATGGAAACTGGATATATGGAGACGTGCTCTGCATAAGCAACCGATATGTGCTTC  409

Qy       339  ACACCAACCTCTACACCAGCATCCTCTTCCTCACTTTCATTAGCATGGACCGATATCTGC  398
              |  ||||||||||| |||||||||  |||||  ||||||||| ||  |||||| || |||||  ||
Db       410  ATGCCAACCTCTATACCAGCATTCTCTTTCTCACTTTTATCAGCATAGATCGATACTTGA  469

Qy       399  TCATGAAGTACCCTTTCCGAGAACACTTTCTACAAAAGAAGGAATTTGCCATTTTAATCT  458
              | ||  |||||  ||||||||||||||| ||||  |||| |||||||||| || ||||||  ||||||||||||
Db       470  TAATTAAGTATCCTTTCCGAGAACACCTTCTGCAAAAGAAAGAGTTTGCTATTTTAATCT  529

Qy       459  CGCTGGCTGTCTGGGCCTTAGTGACCTTAGAAGTTCTACCCATGCTCACTTTCATCAATT  518
              |    ||||  |  |||||   |||||  ||||||||||||||||  |  ||||||||| ||  |  ||  |  ||  |||
Db       530  CCTTGGCCATTTGGGTTTTAGTAACCTTAGAGTTACTACCCATACTTCCCCTTATAAATC  589

Qy       519  CTGTCCCAAAAGAAGAGGGCAGTAACTGCATCGACTATGCAAGTTCTGGAAACCCTGAAC  578
              ||||     ||  ||   | |||| |  |||   ||  |  |||||||||||||| ||||  |
Db       590  CTGTTATAACTGACAATGGCACCACCTGTAATGATTTTGCAAGTTCTGGAGACCCCAACT  649

Qy       579  ACAATCTCATTTACAGCCTCTGCCTGACTTTGTTGGGCTTCCTAATTCCTCTCTCTGTGA  638
              ||||  ||||||||||| |  || || ||  || |||||||  |||||| |||||| ||||||  | |||||
Db       650  ACAACCTCATTTACAGCATGTGTCTAACACTGTTGGGGTTCCTTATTCCTCTTTTTGTGA  709

Qy       639  TGTGCTTCTTCTACTACAAGATGGTAGTCTTCTTAAAGAGGAGGAGCCAGCAGCAAGCAA  698
              ||||  ||||| || ||||||||| |     ||||| ||||| |||||   |||| ||  |
Db       710  TGTGTTTCTTTTATTACAAGATTGCTCTCTTCCTAAAGCAGAGGAATAGGCAGGTTGCTA  769

Qy       699  CTGCCCTGCCACTGGACAAACCCCAACGCCTGGTGGTCCTGGCGGTTGTGATCTTCTCTA  758
              ||||  ||||| ||  ||  ||  || |     | |||| || ||||| ||  || ||||||||||
Db       770  CTGCTCTGCCCCCTTGAAAAGCCTCTCAACTTGGTCATCATGGCAGTGGTAATCTTCTCTG  829

Qy       759  TACTCTTCACACCCTATCATATCATGCGCAATTTGAGGATCGCCTCACGCCTGGATAGTT  818
              | || ||  ||||||||||| |  |||||||  || |||||||||||||||||||||||||  ||||
Db       830  TGCTTTTTACACCCTATCACGTCATGCGGAATGTGAGGATCGCTTCACGCCTGGGGAGTT  889

Qy       819  G---GCCACAAGGATGTACACAGAAGGCCATCAAATCTATATACACACTGACACGGCCTC  875
              |   ||   |   || || ||| | |||||| ||  | ||||    |||||||||||
Db       890  GGAAGCAGTATCAGTGCACTCAGGTCGTCATCAACTCCTTTTACATTGTGACACGGCCTT  949

Qy       876  TGGCCTTTCTGAACAGTGCCATCAATCCCATCTTCTACTTCCTCATGGGAGACCATTACA  935
              |||||||||||||||||||| ||||||  || ||||||| || || ||||||| || | ||
Db       950  TGGCCTTTCTGAACAGTGTCATCAACCCTGTCTTCTATTTTCTTTTGGGAGATCACTTCA  1009

Qy       936  GAGAGATGCTGATTAGTAAGTTCAGACAATACTTCAAGTCCCTTACATCCTTCAGGACAT  995
              | || |||||||| | | |  | ||||| ||||||| |||||||||||||||| || | ||
Db      1010  GGGACATGCTGATGAATCAACTGAGACACAACTTCAAATCCCTTACATCCTTTAGCAGAT  1069

Qy       996  GAGCTGCTGGATGCAGGTCTTCACTCAGCCAAAA-TGAGACACTTGATAAACAG  1048
              | |||  || |  |     |||| |||| |||| ||||     |||| ||||||
Db      1070  GGGCTCATGAACTCCTACTTTCATTCAGAGAAAAGTGAGGGGCTTGTGAAACAG  1123
```

Search completed: August 24, 2004, 18:00:11
Job time : 753 secs